BIOCHEMICAL CHARACTERIZATION AND CRYO-EM STUDIES OF A HIGHLY ACTIVE SPINACH PHOTOSYSTEM II COMPLEX

A Dissertation
Presented to
The Academic Faculty

by

Yusuf M. Uddin

In Partial Fulfillment
of the Requirements for the Degree
Ph.D. in the
School of Biological Sciences

Georgia Institute of Technology
December 2018

BIOCHEMICAL CHARACTERIZATION AND CRYO-EM STUDIES OF A HIGHLY ACTIVE SPINACH PHOTOSYSTEM II COMPLEX

Approved by:

Dr. Ingeborg Schmidt-Krey, Advisor
School of Biological Sciences
Georgia Institute of Technology

Dr. Loren Williams
School of Chemistry & Biochemistry
Georgia Institute of Technology

Dr. Bridgette Barry
School of Chemistry & Biochemistry
Georgia Institute of Technology

Dr. Nael McCarty
Emory + Children's Center for Cystic Fibrosis and Airways Disease Research
Emory University

Dr. Al Merrill
School of Biological Sciences
Georgia Institute of Technology

Date Approved: November 1, 2018

ACKNOWLEDGEMENTS

This project and thesis is a reflection of the hard work and dedication towards a task that seemed very impossible. The far-reaching goal set in the early years of my program would not have been possible without the support of my advisor, Dr. Ingeborg Schmidt-Krey, and my collaborating advisor, Dr. Bridgette Barry, who were both gracious enough to allow me to work with them. They have been my mentors and given me guidance and support throughout these years.

I would like to thank the current and former members of the Barry lab (Udita Bramachari, Dr. Brandon Polander, Dr. Zhanjun Guo, Jiayuan (Nancy) He, Chichi Obi, and Sara Konecny, Dr. Jiafeng Geng, Dr. Krista Cosert, Dr. Atlee Watson, and Tyler McCaslin) who were instrumental in teaching me the purification, biochemical characterization, and helpful discussions. Thank you to Dr. Smalley for his assistance on the proteomics work at Georgia Tech's Systems Mass Spectrometry Core Facility.

On the cryo-EM and image processing work, Drs. Jay Rai & Mykailo Kopylov (Stroupe Lab, Florida State University), Kasahun Neselu, Michael Sofroniou, Eric Woods, and Carolann Espy all provided helpful assistance. Thank you to the staff at Georgia Tech's Partnership for an Advanced Computing Evironment (PACE). I would also like to thank current and former members of the Schmidt-Krey lab (Dr. Matthew C. Johnson, Kerry Strickland, Maureen Metcalfe, and Justin Dehorty) for their helpful discussions pertaining to electron microscopy and this project. The collaboration with the Stroupe lab and the staff at the Biological Science Imaging Resource (Dr. Nilakshee Battacharya, Dr. Duncan Sousa, and Dr. Xiaofeng Fu) at Florida State University via the

NIH U24 consortium was instrumental to this work, and we appreciate the help they have given us. Finally, I would like to thank my thesis committee, the School of Biological Sciences, and the Parker H. Petit Institute for Bioengineering and Biosciences.

On the personal end, I want to thank my parents, Mohammed and Chong Uddin, my sisters, Maya, Sophia, and Sonya, and all my friends who have supported me through this tough journey. The accomplishments that I have made would not have been possible without all my colleagues, friends, and family, and if you are reading this dissertation, thank you for your time and interest.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

2D	two-dimensional
3D	three-dimensional
Å	Ångstrom
BBY	Berthold, Babcock, Yocum preparation
C_2	core 2
C_2S_2	core 2 strongly bound 2
$C_2S_2M_2$	core 2 strongly bound 2 moderately bound 2
CCD	charged-coupled device
CCF	cross correlation function
Chl	chlorophyll
CMOS	complementary metal-oxide semiconductor
Cryo-EM	electron cryo-microscopy
DDM	dodecyl β-D-maltoside
DGDG	diagalactosyldiacyl glycerol

DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
FFT	fast Fouier transform
FSC	Fourier shell correlation
HPLC	high performance liquid chromatography
kDa	kiloDalton
LPR	lipid to protein ratio
MES	2-(*N*-morpholino)ethanesulfonic acid
mg	milligram
MGDG	monogalactosyldiacyl glycerol
mL	milliliter
NSAF	normalized spectral abundance factor
OEC	oxygen evolving complex
OTG	octyl β-D-thioglucopyranoside
OTG-PSII	OTG-photosystem II
OPDM	octyl thioglucopyranoside photosystem II in dodecyl maltoside
PDB	Protein Data Bank

PC	phosphatidyl choline
PG	phosphatidyl glycerol
PSII	photosystem II
PSM	peptide spectral match
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamid gel electrophoresis
SNR	signal to noise ratio
SPA	single particle analysis
SQDG	sulphoquinovosyldiacylglyceerol
TEM	transmission electron microscopy
TX-100	Triton X-100
UV-vis	ultraviolet-visible spectroscopy

SUMMARY

Photosystem II (PSII) is a large, photosynthetic membrane protein complex responsible for water oxidation and the formation of oxygen on Earth. Understanding the structure of this protein is important for studying photosynthetic energy transfer and assembly mechanisms in higher plants. Recently, technological breakthroughs in the field of cryo-EM have lead to the ability to study this membrane protein by single particle analysis (SPA). This technique does not require the growth of crystals, a challenge that has hindered structural determination in higher plant PSII. In this work, the biochemical preparation of spinach PSII for cryo-EM is optimized to maintain high activity. Subunit characterization, oxygen activity, UV absorbance spectroscopy, mass spectrometry, and negative stain transmission electron microscopy were employed to characterize the spinach PSII preparation. Cryo-EM experiments resulted in a first model of a highly active C_2 PSII complex, which contains all intrinsic and extrinsic subunits essential for activity.

CHAPTER 1. INTRODUCTION

1.1 Photosystem II: A Multi-Subunit Photosynthetic Enzyme

A few billion years ago, the Earth was not as rich with oxygen as it is today. Life relied on resources from the environment, such as hot springs and other geothermal sources [1, 2]. Eventually, prokaryotic cells evolved a way to utilize energy from light which was orders of magnitude more efficient. This process is known as photosynthesis and is the reason for the abundance of oxygen in the atmosphere. This light driven reaction is performed in cyanobacteria, algae and plants [3]. At the heart of photosynthesis is a membrane protein critical in the conversion of water into molecular oxygen: photosystem II.

1.1.1 Background

PSII is found in the thylakoid membranes of cyanobacteria and in the cells of plants and algae (Figure 1.1). Aside from PSII, other membrane proteins involved in the photosynthetic electron transport chain are cytochrome b_6f, photosystem I, and ATP synthase. The ensemble of proteins work together to harvest light energy and convert water into oxygen and protons, while simultaneously creating ATP and NADPH, which are used as chemical energy and reducing agents necessary for CO_2 fixation [4]. The majority of PSII is found within the grana stacks, while the other proteins are found near the stromal-facing edges, or grana margins [5, 6].

The light reactions of photosynthesis make up the basis of the Z-scheme, part of which photosystem II plays the first major role [7]. The mechanism of PSII (Figure 1.2)

starts with a photon that excites the cluster of chlorophyll molecules called P680 at the center of the complex [8-10]. This ejects an electron from the pigment cluster, which travels to a pheophytin molecule and then reduces Q_A and creates a plastoquinol [11]. The oxygen evolving center (OEC), which is a Mn_4CaO_5 cluster, is oxidized by a nearby amino acid residue called tyrosine Z, which replenishes the electron lost from $P680^+$. This process happens 4 times before molecular oxygen is created. Equation 1.1 show a summary of the photosynthetic reaction.

$$2H_2O + light \rightarrow O_2 + 4H^+ + 4e^- \qquad \textbf{Equation 1.1}$$

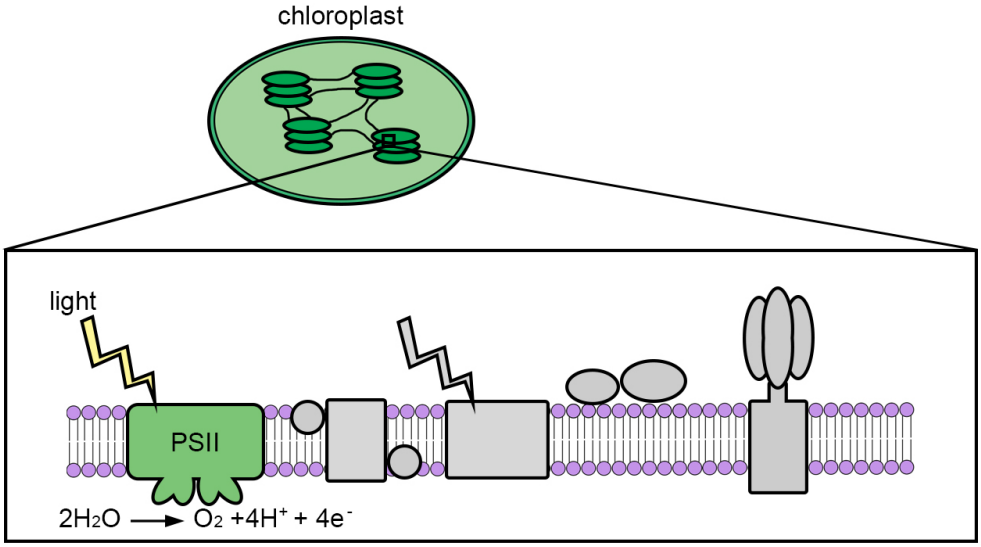


Figure 1.1 Photosynthetic electron transport chain. Plant PSII is found in the thylakoid membranes of chloroplast. Light powers the complex to oxidize water into molecular oxygen and hydrogen. The electrons are passed to the other complexes in the electron transfer pathway.

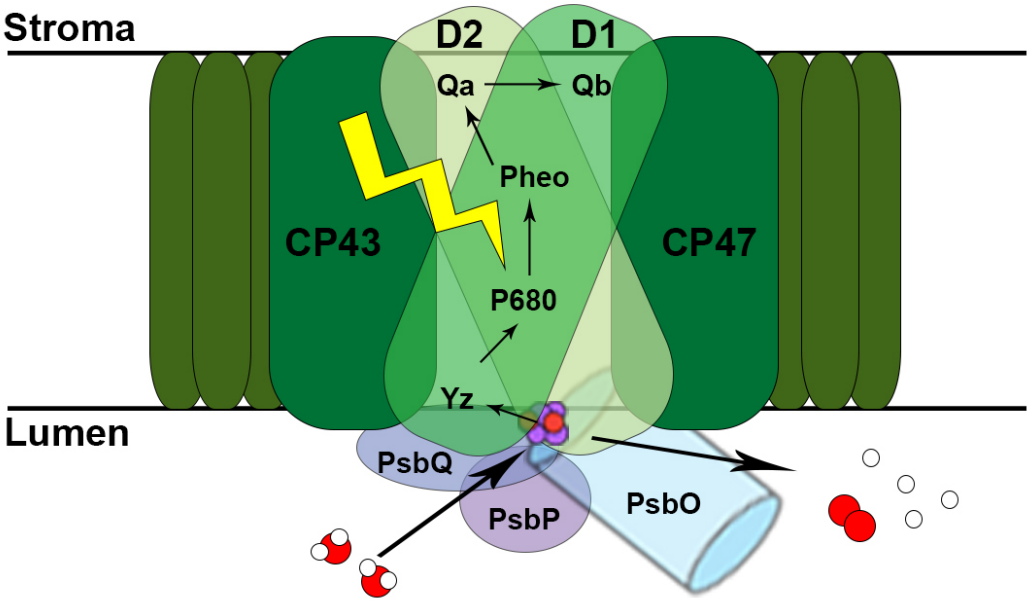


Figure 1.2 General mechanism of electron transport in PSII. Located within D1 and D2 polypeptides are most of the components involved in PSII electron transport. Light is funneled down to P680 which ejects an electron to pheophytin. Water near the Mn_4CaO_5 cluster is oxidized to replenish the electrons lost, ultimately forming O_2 and a proton gradient after 2 complete cycles.

It is known that the OEC undergoes a cyclic pattern of oxygen evolution. More specifically, the active site can be flashed with four short bursts of light which will transition the active site through a series of four S states. S_0-S_4 states describe the manganese cluster at different points in time where electrons and protons are released (Figure 1.3). Together, this repetitious pattern makes up the Kok cycle [12, 13]. After one complete cycle, the extraction of four electrons and four protons from H_2O occurs, releasing O_2 into the lumen. The manganese cluster acts as a charge storage device which passes electrons to $P680^+$, which continue to pass them through the electron transport chain. When the electrons arrive to Q_B, they are combined with two protons in the stroma

to create a non-polar, plastoquinol (PQH_2), which can migrate through the membrane to cytochrome b_6f [12].

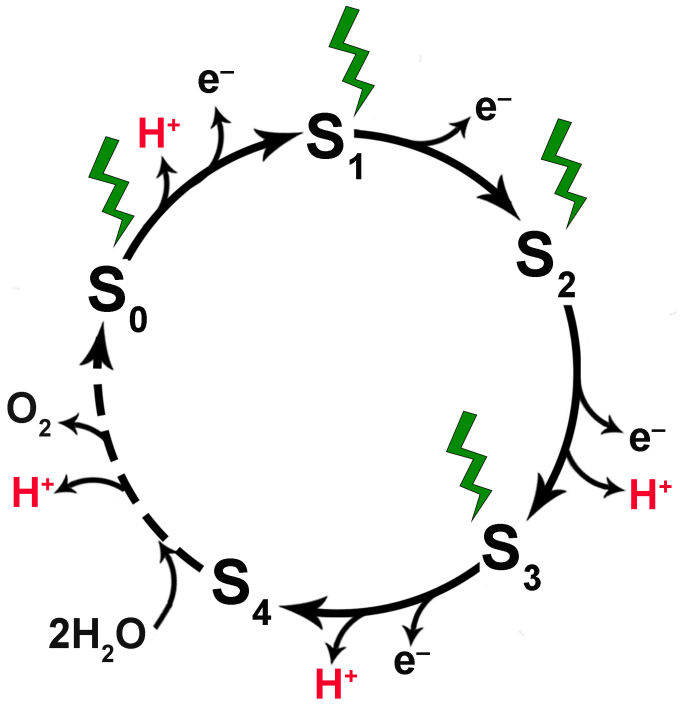


Figure 1.3 Kok cycle describing the water splitting reaction in PSII [12].

1.1.2 Structural information of PSII aids in understanding mechanism

A detailed map of the subunits and cofactors that are involved in the PSII reaction is necessary when studying the mechanism. A high resolution structure of PSII was solved in 2011 which gave the first insights to the structure of the OEC and the proposed mechanism of electron transport [14]. This structure was solved from a thermophilic cyanobacteria, *Thermosynechococcus vulcanus*, which has a well-characterized crystallization method. Unfortunately for higher plants like PSII, 2D and 3D

crystallization methods have been attempted, but none have resulted in high resolution structures [15-17]. Though there are many conserved subunits between the structures of cyanobacteria and higher plant PSII, there are still a large number of differences. In 2016 and 2017, high resolution cryo-EM structures of plant PSII were solved. However, these PSII samples were not highly active [18-20]. As with crystallography, it is preferable to maintain homogeneity during sample purification for SPA cryo-EM. This can sometimes sacrifice enzymatic activity, which is possibly what occurred in the recent cryo-EM structures. Nonetheless, the protein structures provide important details about the organization of subunits, which answered questions about the evolutionary differences across species. It also located the positioning of chlorophylls in the peripheral antenna proteins. Knowledge of these locations helped us understand how energy is funneled from the light harvesting proteins down to the core of the reaction center.

1.1.3 Lipids found in the thylakoid membranes

The thylakoid membranes contain a mixture of lipids. These lipids provide a matrix for the complexes involved in the photosynthetic electron transport chain [21]. They maintain the structural integrity of the thylakoid membrane to regulate the movement of important molecules and also allow flexibility in protein complex repair. The membranes are primarily composed of monogalactosyldiacyl glycerol (MGDG), diagalactosyldiacyl glycerol (DGDG), sulphoquinovosyldiacylglycerol (SQDG), phosphatidyl glycerol (PG) and phosphatidyl choline (PC) [22]. Even upon purification of PSII, some of these lipids are found intercalated between subunits, which indicates that they not only maintain structure, but also have a functional role in the development and processing of the complex [23]. Among the core intrinsic subunits, there are 4 MGDG

lipids, 3 DGDG lipids, 3 SQDG lipids, and 3 PG lipids (Table 1.1) [18]. Lipids play an important role in the regulation of oxygen evolution. Studies have shown that the addition of exogenous lipids can increase the levels of oxygen evolution, and that adding acidic lipids alters the overall organization of the PSII structure and the pH environment [24].

1.1.4 Pigment molecules of PSII

Chlorophyll is an essential pigment involved in photosynthesis. PSII is primarily composed of both chlorophyll *a* and *b* pigments. Chlorophyll *a* differs from chlorophyll *b* at the C-7 position of the chlorin head group. Chlorophyll *a* has a methyl group at that position, and chlorophyll *b* has an aldehyde. Beyond this difference, the structures of the pigments are identical. However, chlorophyll *a* absorb mostly violet and red wavelengths, specifically near 440 and 660 nm, while reflecting green light [25]. In order to broaden the spectrum of light absorbed for photosynthesis, chlorophyll *b* and other carotenoids are synthesized in the plant. Their function is to absorb and transfer photons to the chlorophyll *a* dimer, P680, in order for the PSII reaction to occur [26].

In higher plants, the light harvesting antennae proteins contain higher quantities of chlorophyll *b,* while the major core proteins of PSII contain an abundance of chlorophyll *a*. In addition to chlorophyll pigments, carotenoids (β-carotene, lutein, neoxanthin, and violaxanthin) are also found scattered through the subunits (Table 1.1) [18].

Table 1.1 Cofactors found in subunits of plant PSII [18, 19].

Protein	Lipid	Chlorophyll	Carotenoid	Others
D1	2 SQDG 1 MGDG	4 Chl *a* 2 Pheophytin	1 β-carotene	1 Mn_4CaO_5 cluster 1 plastoquinone
D2	3 PG 1 MGDG	2 Chl *a*	1 β-carotene	1 plastoquinone
CP43	3 DGDG 1 MGDG	13 Chl *a*	3 β-carotene	
CP47	1 SQDG 1 MGDG	16 Chl *a*	3 β-carotene	
LHCII trimer (S and M)	3 PG	24 Chl *a* 18 Chl *b*	6 luteins 3 violaxanthins 3 neoxanthins	
CP26	1 PG	9 Chl *a* 4 Chl *b*	2 luteins 1 neoxanthins	
CP29	1 PG	10 Chl *a* 3 Chl *b*	1 luteins 1 violaxanthins 1 neoxanthins	
CP24	1 PG	6 Chl *a* 5 Chl *b*	1 luteins 1 violaxanthins 1 β-carotene	

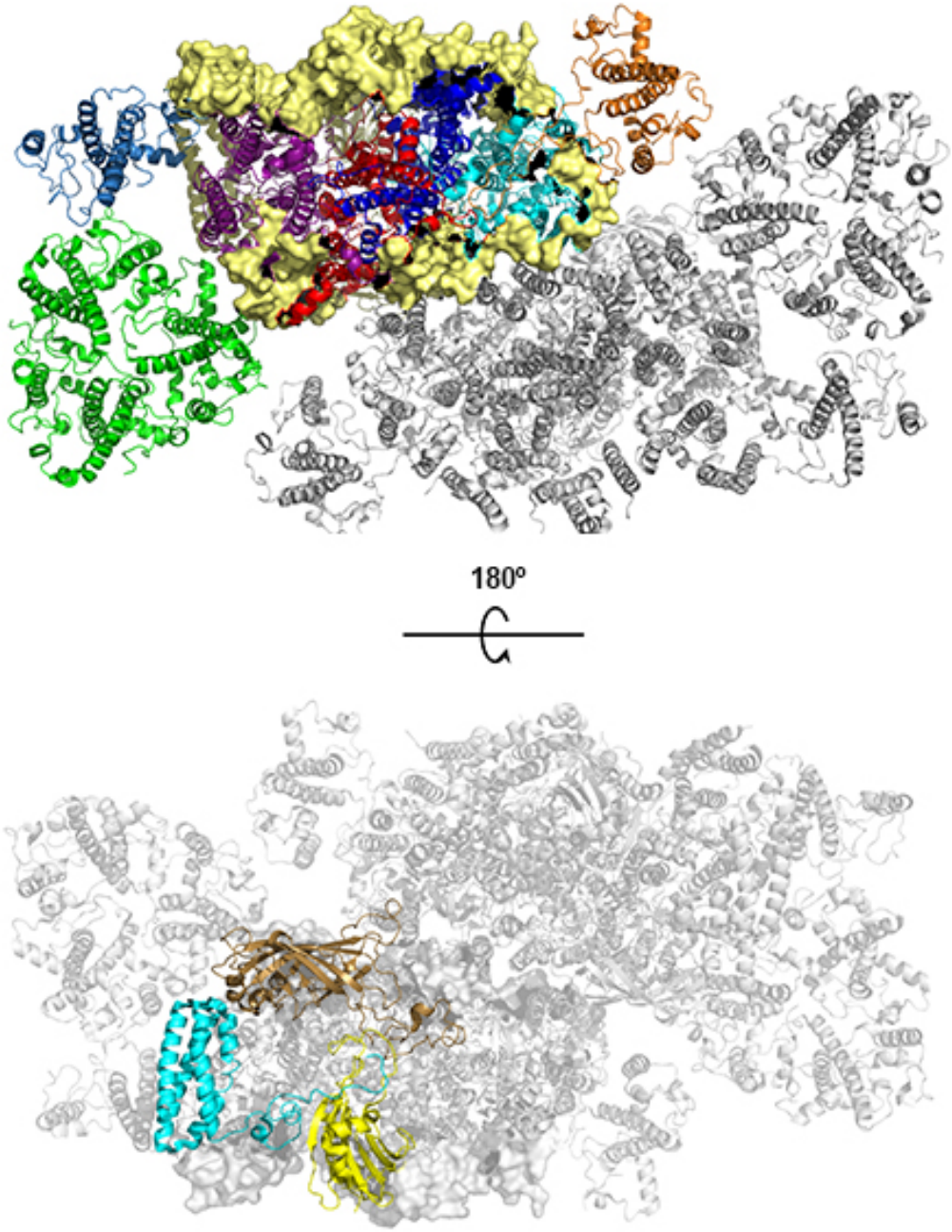


Figure 1.4 Structure of plant PSII. Top structure is of the stromal-facing side of PSII, and bottom structure is of the luminal-facing side. D1 (red), D2 (blue), CP47 (teal), CP43 (purple), are shown on one monomer. Antenna proteins, LHCII timer (green), CP26 (light blue), and CP29 (orange) are also shown. The small intrinsic subunits (light yellow) are shown in surface view. The extrinsic subunits are PsbO (brown), PsbP (yellow), and PsbQ (cyan). Structure derived from PDB 3JCU [18].

1.1.5 The subunits of plant photosystem II

Depending on the organism, there are approximately 20 subunits that make up PSII, along with cofactors like pigment molecules, metals, and lipids [21]. These proteins make up the entire PSII complex and can exist in many different quaternary arrangements due to the complex's dynamic nature. The evolutionarily conserved "core" intrinsic subunits (CP47, CP43, D1, D2) and extrinsic, peripheral membrane protein subunits (PsbO, PsbP, and PsbQ), are the focus of this work. The core integral membrane proteins are found across all photosynthetic species, but the extrinsic subunits vary in composition depending on the species. Thus, it is important in our plant PSII preparation that these subunits are included, intact, and properly oriented in order for a highly active structure of PSII to be solved. Each of the major and minor subunits and will be discussed briefly in this section, using Figure 1.4. Additionally, Table 1.2 provides some details such as gene names, molecular weights and basic functions.

1.1.5.1 CP43 and CP47

The CP43 and CP47 polypeptides are involved in energy transfer and hence they are sometimes called the core antenna proteins (Figure 1.4, purple and teal, respectively). They contain six transmembrane α-helices with several loops that extend out of the stroma and lumen. Despite the lack of sequence homology, the two proteins look and are arranged in a similar way. They have 11 conserved histidine residues, however, which is thought to play an important role in binding chlorophyll *a*. Previous studies have genetically mutated these proteins to remove the histidine amino acids and the results showed deplete light harvesting capabilities [27]. They also aid in the binding and

protection of the extrinsic subunit, PsbO, which is important for oxygen activity. Together, the chlorophyll abundance and carotenoids allow the CP43 and CP47 subunits to funnel the excitation energy to the reaction center [28, 29].

1.1.5.2 D1 and D2

D1 and D2 subunits contain the important ligands involved in electron transfer (Figure 1.4, red and blue, respectively). They contain 5 transmembrane helices as well as several important loops and small helices in the stroma and lumen [30]. Nearby are two α -helices called cytochrome b559, which bind heme and facilitate in photo-protection. These subunits house P680, the primary electron donor. D1 contains a tyrosine residue (Y_Z) that aids the the movement of electrons from the OEC to P680. A second redox active tyrosine residue (Y_D) is found in D2 in a symmetrical location to Y_Z. D1 and D2 subunits are evolutionarily conserved across photosynthetic species and play a structural and functional role similar to that of the purple bacteria reaction centers [31]. Even though PSII requires light to function, it can also lead to a type of damage called photoinhibition, specifically to the D1 protein. There is a constant turnover of this subunit, which requires degradation of the damaged protein, partial disassembly of PSII, movement across the thylakoid membrane, and replacement with a new D1 subunit [32-34]. The extent of this photoinhibition results from a balance of the rates between photo-damage and photo-repair of these subunits [35].

1.1.5.3 Oxygen evolving complex

The OEC is not a protein subunit, but rather an inorganic Mn_4CaO_5 cofactor embedded in the D1 and D2 subunits. Its presence in PSII is critical for activity of the

complex. The OEC acts as an electrical accumulator, by being oxidized by the $P680^+$ four times before replenishing its electrons from water, creating dioxygen, and resetting itself back to its original, most reduced state [36]. The overall "molecular picture" of each of these S states is only halfway solved, due to spectroscopic and structural data reporting that only the S_1 and S_2 states exist in the current data [37]. Regardless of the states, the cluster adopts a cubane-like structure resembling a distorted chair with oxygen atoms connecting each of the metals to one another (Figure 1.5) [14]. There are several chloride ions that also exist near the cluster that aid in stabilization. Loss or removal of the extrinsic proteins destabilizes the OEC and reduces overall oxygen evolving activity. This is also why $CaCl_2$ restores oxygen activity in purifications where the extrinsic subunits have been lost [38].

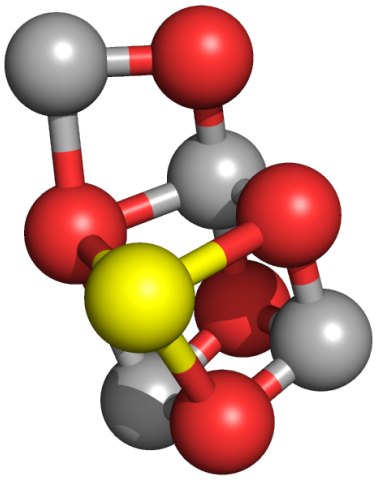


Figure 1.5 Model of the oxygen evolving complex in PSII. The Mn_4CaO_5 cluster is found embedded in each monomer of active PSII. Grey spheres are manganese, red spheres are oxygen, and calcium is in yellow. Model derived from PDB 4UB6 [39].

1.1.5.4 Extrinsic polypeptides

Found on the luminal side of PSII, the extrinsic proteins play a big role in maintaining activity of the complex (Figure 1.4, PsbO brown, PsbP yellow, PsbQ cyan). PsbO, sometimes called the manganese stabilizing protein or the 33 kDa protein, is a larger, β-barrel subunit that is evolutionarily conserved across photosynthetic species. The presence or absence of PsbO does not affect photoautotrophic growth in cyanobacteria, but alga and plants require it [40]. In solution, it is known to be an intrinsically disordered protein, but when bound to PSII, it adopts its β-barrel structure [41]. PsbO is involved in stability of the OEC and binding of calcium and chloride. Loss of PsbO causes a large drop in the overall activity of PSII. There is substantial biochemical evidence that two copies of PsbO per monomer of PSII exist [42-44]. However, most PSII structures have only indicated electron density for one copy per monomer. It could be possible that the purification methods may be too harsh for the complex to retain all copies of the extrinsic subunits.

PsbP and PsbQ, (also known as the 24 and 18 kDa subunits, respectively) both have roles in calcium and chloride binding as well, though their removal from the complex is not as detrimental, provided that there is a high concentration of calcium and chloride ions to stabilize the OEC [43]. In cyanobacteria, PsbU and PsbV are the protein equivalents to the plant PsbP and PsbQ. In red and green algae, the presence of cyanoP and cyanoQ extrinsic subunits are found, as well as PsbU and PsbV.

Together in plant photosystem II, these three extrinsic proteins form a crown-like structure on each monomer of PSII. They are bound to to the CP43 and D1 intrinsic

subunits. The recent cryo-EM structure revealed a nuclear encoded PsbTn subunit with unknown function [18].

It has been considered that the placements of these extrinsic subunits may still be in question, based on several previous works which conflict with the recent positioning of PsbO, PsbP and PsbQ. These studies remove the extrinsic subunits, PsbP and PsbQ by performing a salt wash treatment and subsequent treatment to remove all extrinsic subunits [42, 45, 46]. The structural data, granted its lower resolutions, indicate that there might be an alternate binding, where PsbO and PsbP bind at opposite locations. According to cyanobacteria structures, the placement of PsbO is the same as the recent higher plant structures, which were derived from cyanobacterial models of PSII.

1.1.5.5 Low molecular weight intrinsic polypeptides

Forming a belt around the core complex (Figure 1.4, light yellow), are 12 low molecular weight subunits (PsbE, F, H, I, J, K, L, M, Tc, W, X, Z), usually containing 1 transmembrane helix (Table 1.2). The majority of these small proteins have structural homologues in cyanobacteria. About half of these subunits have their N-termini facing the stroma, and the other half have the N-termini facing the lumen [18]. Depending on which side of the core monomer they are on, their roles are to provide structural support, dimer stability, stabilization of the core, and maintain association with the peripheral antennae proteins [47]. The presence of these small intrinsic polypeptides can be thought of as the molecular glue holding the PSII complex together.

1.1.5.6 Light harvesting antennae subunits

Along the periphery of the plant PSII core dimer, the antennae proteins are docked in varying combinations (Figure 1.4, LHCII green, CP26 light blue, CP29 orange). The traffic of these proteins varies depending on the environment. Cyanobacteria and red algae utilize structures called phycobilisomes, which lie outside of the membrane and absorb light energy in a wavelength region of 500 and 650 nm [48, 49]. In plants, there are 6 *lhcb* genes that encode proteins that contain 3 main α-helices, and adopt similar structural patterns. The first three genes produce a trimer called LHCII (light harvesting complex). This complex defines the nomenclature for the strongly bound LHCII and moderately bound LHCII versions of the PSII supercomplex [50]. The other three *lhcb* genes produce CP24, CP26, and CP29. The CP26 and CP29 monomers are found in the C_2S_2-type PSII supercomplex, while CP24 is found in the $C_2S_2M_2$ supercomplex. These antennae proteins contain both chlorophyll *a* and *b*, but are rich in chlorophyll *b*. This is important because these peripheral proteins are responsible for collecting photons and passing it down the energy funnel to eventually reach the P680 molecules [19, 20, 51].

Table 1.2 Major and minor subunits of plant photosystem II [18, 47].

Gene	Subunit	Mass (kDa)	Function
psbA	D1	38	Contains tyrosine Z residue. Binds P680, pheophytin, Qb.
psbB	CP47	56	Excitation energy transfer, binds PsbO
psbC	CP43	50	Excitation energy transfer, binds PsbO
psbD	D2	39	Contains tyrosine D residue. Binds P680, pheophytin, Qa.
psbE	α-cyt b559	9	Binds heme, core stability, photoprotection
psbF	β-cyt b559	4	Binds heme, core stability, photoprotection
psbH	PsbH	8	Aids in binding antennae complexes
psbI	PsbI	4	Dimer stability
psbJ	PsbJ	4	Core stability

Table 1.2 continued.

Gene	Subunit	Mass (kDa)	Function
psbK	PsbK	4	Core stability
psbL	PsbL	4	Dimer stability
psbM	PsbM	4	Dimer stability
psbO	PsbO	26 (33 in gel)	Stabilizes OEC, Ca^{2+} and Cl^- binding
psbP	PsbP	20 (24 in gel)	Ca^{2+} and Cl^- binding
psbQ	PsbQ	16 (18 in gel)	Ca^{2+} and Cl^- binding
psbR	PsbR	10	Donor and acceptor side functions
psbS	PsbS	22	Chlorophyll chaperonin
psbT	PsbTc/PsbTn	3	Nuclear encoded variant: Bridges CP47 and PsbE. Chloroplast encoded variant: Dimer stability
psbW	PsbW	6	Aids in binding antennae complexes

Table 1.2 continued.

Gene	Subunit	Mass (kDa)	Function
psbX	PsbX	4	Core stability
psbZ	PsbZ	7	Aids in binding antennae complexes. Contains 2 α-helices
lhcb1	Lhcb1	25	Light harvesting, chlorophyll b rich
lhcb2	Lhcb2	25	Light harvesting, chlorophyll b rich
lhcb3	Lhcb3	25	Light harvesting, chlorophyll b rich
lhcb4	CP29	29	Excitation energy transfer & dissipation
lhcb5	CP26	26	Excitation energy transfer & dissipation
lhcb6	CP24	24	Excitation energy transfer & dissipation

1.1.6 Quaternary assembly of PSII

PSII exists as a dimer *in vivo*, though depending on the purification, there are many variants of complexes that form [52]. Though each type is still considered PSII, a nomenclature has been given to distinguish smaller types from larger types [50]. As stated before, in a response to the amount of light present in the environment, plants will arrange the content of their thylakoids to either accept more light, or reduce the amount. During times where there is little light present, PSII will have many accessory antenna proteins that help funnel more light to its active center [51]. In Figure 1.6, three examples are given of possible assemblies of PSII that are found when purifying the protein complex by sucrose gradient ultracentrifugation. C_2 is the core complex, consisting of the major and minor intrinsic proteins and the extrinsic subunits. C_2S_2 supercomplex, sometimes called the C_2S_2-LHCII, contains the antennae proteins: CP26, CP29, and LHCII strongly-bound trimer, in addition to the C_2 core components. $C_2S_2M_2$ refers to the megacomplex with the additional antennae protein, CP24, and a moderately bound LHCII trimer. In the sucrose gradient purification of PSII, it is possible to find varying combinations of each of these three types as well as asymmetrical versions [50].

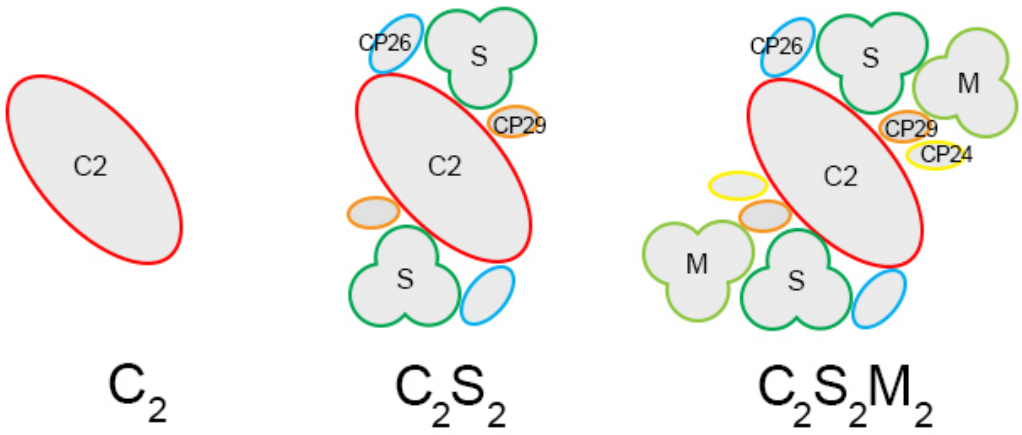


Figure 1.6 Quaternary arrangements of the plant PSII complex.

1.1.7 Differences in PSII structures across species

The most obvious difference between cyanobacteria PSII and higher plant PSII, is the evolution of intrinsic antennae subunits in higher plants compared to the large phycobilisomes of cyanobacteria. Additionally, the extrinsic subunits of cyanobacteria are PsbO, PsbU, and PsbV, while higher plants have PsbO, PsbP, and PsbQ proteins. Although, as one proceeds higher up on the evolutionary timeline, one begins to see the emergence of cyanoP and cyanoQ extrinsic subunits in the algae species [53]. As stated earlier, the main intrinsic subunits of PSII are highly conserved (CP43, CP47, D1, D2) across species such as cyanobacteria, euglena, red and green algae, and higher plants [54-57]. The reaction center of PSII is highly stable across photosynthetic species largely due to the evolutionary pressures on light harvesting and photo-protection [58, 59].

Very recently, there have been five cryo-EM structures of plant PSII, all of which were solved of PSII with compromised oxygen evolving activity [18-20, 60]. Table 1.3 describes the recent structures of plant PSII. The two types of complexes were the C_2S_2 and $C_2S_2M_2$. It is important to note that all five structures were purified using sucrose gradient centrifugation, which causes a decline in oxygen evolving activity and can lead to the loss of important extrinsic subunits. Additionally, some purifications protocols use buffers with pH 7.5 which causes a loss of activity. As outlined earlier, a uniform, homogenous protein complex sample, rather than a heterogeneous mixture, greatly facilitates single particle analysis. For these structures, the cost of subjecting the sample to a polishing step such as sucrose gradient or gel filtration was worth the effort in obtaining the initial reconstructions at high resolution. However structural studies of a highly active PSII complex are critical to fully elucidating the detailed reaction mechanism.

Table 1.3 Recent cryo-EM plant PSII structures.

PDB code	Year	PSII complex size	Plant Species	Resolution	O_2 Evolving Activity ($\mu mol\ O_2\ mg\ chl^{-1}\ hr^{-1}$)
3JCU	2016	C_2S_2	*Spinacia olerasea*	3.2 Å	75
5XNN	2017	$C_2S_2M_2$ (stacked)	*Pisum sativum*	2.7 Å	288
5XNM	2017	$C_2S_2M_2$ (unstacked)	*Pisum sativum*	3.2 Å	117
5MDX	2017	$C_2S_2M_2$	*Arabidopsis thaliana*	5.3 Å	not reported
N/A	2017	C_2S_2M (stacked)	*Pisum sativum*	14 Å	not reported

1.1.8 Maintaining activity of PSII is critical for understanding mechanism

A high-resolution structure of fully active PSII is important when understanding the mechanism of electron transport. Altered activity of PSII is usually indicative of damage to the structure of the OEC, loss of important extrinsic subunits, or conformational changes across the entire complex. Lowered rates in plant PSII have been associated with the loss of extrinsic polypeptides, PsbO, PsbP, and PsbQ, but current structural data do not relate these subunits to direct locations of cofactors [61]. Even the loss of just PsbP and PsbQ can result in reduced oxygen evolution [41].

The method used for structural elucidation is an important consideration when understanding the integrity of the protein. It has been reported that the PSII crystal sustained damage to the redox-active metal cluster during XRD data collection [62]. This damage, determined by spectroscopic methods, reduces the manganese to different electronic states, which effects the overall structure of the OEC by increasing ligand distances [62-64]. A paper on radiation-free femtosecond X-ray structures of undamaged PSII complex in cyanobacteria, explains that there are significant differences between the arrangement of the OEC compared to the structure of PSII elucidated by synchrotron methods [39]. Therefore, if such important changes occur in the X-ray structures in cyanobacteria PSII, the differences between an active and inactive PSII structure must yield crucial information.

1.2 Current Challenges of Membrane Proteins in Structure Determination

The discovery and utilization of X-ray diffraction in the 19th and early 20th centuries eventually paved the way for protein structure elucidation [65]. X-ray crystallography is an important tool for elucidating the structural and functional details of proteins, however many of the structures deposited into the Protein Data Bank (PDB) are of soluble proteins. When analyzing most organism proteomes, 20-30% of the naturally occurring proteins are membrane proteins, and among them, more than 40% are drug targets [66, 67]. Membrane proteins are difficult to purify and crystallize (outlined below), hence they do not make up the majority of structures deposited in the PDB. Despite challenges in structural elucidation of membrane proteins, high resolution information is important in understanding enzymatic mechanisms, drug design, and biotechnology [68, 69].

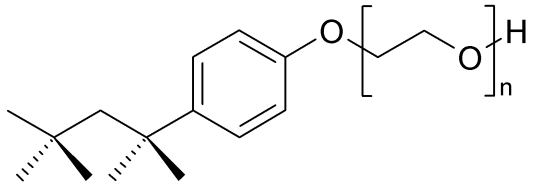
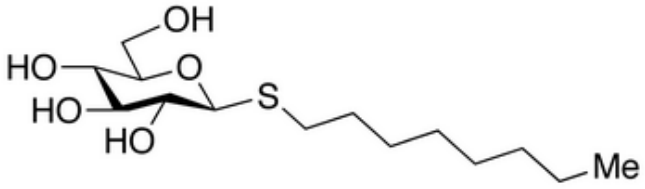
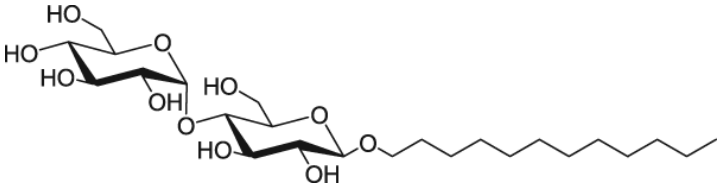
1.2.1 Purification of membrane proteins using detergents

Membrane proteins have hydrophobic and hydrophilic regions because they are located partially inside the membrane bilayer and partially in the solvent. In addition, they pose challenges in expression, purification, and solubilization. Membrane proteins with multiple subunits may lose critical components during purification. Therefore, selecting the right detergents is critical in removing them from the membrane and solubilizing them [70-72]. In some cases, overexpression of a membrane protein is necessary to increase the purification yield. If the protein source is naturally rich within a cell, it is possible to extract the protein without overexpression through a host organism. Membrane proteins are generally stable while they are inside the membrane, but once the

cell is broken, precautions must be taken to avoid protease digestion [72]. Protease inhibitors, maintaining low temperatures, and stabilizing buffered solutions all aid in a resulting active, purified protein. Peripheral membrane proteins can simply be removed from the membrane using high salt or high pH washing, however, for integral membrane proteins, an amphipathic molecule must be utilized [73].

Detergents fall into ionic, nonionic, and zwitterionic classes. Typically, nonionic detergents are useful in removing the membrane protein from the lipid bilayer. The optimal concentration is determined empirically by varying the concentrations, temperature, and solubilization times. Table 1.4 lists the structures and qualities of detergents used throughout the PSII purification. When a membrane protein is found in the supernatant after a centrifugation with detergent, it is generally considered solubilized [72, 74, 75]. Detergents such as Triton X-100, sodium deoxycholate, CHAPS, and octylglucoside are a few common detergents used to purify membrane proteins and maintain activity. Sometimes it may be necessary to use a combination of detergents to successfully purify the membrane protein of interest. Detergents such as dodecyl maltoside (DDM), Brij-35, and digitonin have been successfully used for solubilizing membrane proteins for TEM imaging [76]. A key characteristic that plays a role in the selection of detergents is the critical micelle concentration (CMC). The CMC is the concentration at which the detergent molecule exists as a micelle, and it is necessary to be at or above the CMC for detergents to solubilize membrane proteins [73].

Table 1.4 Structures of detergents used for PSII purification.

Structure	Detergent
	**Triton X-100** MW: 647 Da CMC: 0.24 mM
	**Octyl β-D-thioglucopyranoside** MW: 308 Da CMC: 9 mM
	**Dodecyl β-D-maltoside** MW: 511 Da CMC: 0.15 mM

1.3 Electron Cryo-Microscopy

X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and electron cryo-microscopy (cryo-EM), and are common methods used to elucidate structure. The most frequently used technique has been X-ray crystallography. NMR is a useful technique for smaller complexes (with a molecular weight around 40-60 kDa).

However, both of these methods require a relatively large amount of purified sample [77]. Cryo-EM, utilizes significantly less sample and can be used for the structure determination of heterogeneous samples. In addition, recent advances have lead to significant increases in resolution and structure determination of even smaller samples [78].

1.3.1 Transmission electron microscopy

Transmission electron microscopy (TEM) is analogous to light microscopy in many ways. Instead of light, an electron beam is generated from a tungsten or lanthanum hexaboride filament, or a field emission gun. A series of electromagnetic lenses are used to focus the accelerated electrons through the column of the microscope. Because the wavelength of electrons is around 2.5 picometers, images can be formed of specimens at sub-nanometer to near-atomic level resolutions [79]. One caveat to using electrons is the high potential of scattering from particles in the air. Therefore, an electron beam must be contained in a vacuum environment [80]. For biological specimens, this means precautionary measures must be taken in order to protect the sample from radiation damage due to the electron dose and evaporation due to the vacuum. As the electrons pass through the specimen, the perturbed elastic and inelastic scattered electrons are focused by more electromagnetic lenses and detected, usually by a type of camera detector. Electrons that pass through a three dimensional specimen produce a two dimensional projection [81, 82].

When understanding how 3D reconstructions can be obtained from 2D images, it is important to turn to the Central Slice Theorem (projection theorem) of Fourier transforms

[82, 83]. A central slice, or central section, is a 2D plane that goes through the origin of a 3D Fourier transform. When a 2D projection of an object is created in a transmission electron microscope, the Fourier transform of that real-space projection contains the same information as its central slice in reciprocal space. Therefore, with enough 2D projections at various rotational angles, one can merge the data to create a complete 3D Fourier transform of the object, which can be reverse Fourier transformed to generate the original volume in real-space [84]. The Central Slice Theorem assumes two things: the sample should be thin enough to ignore depth of field so that a single defocus value or range is used. It also assumes the sample is a weak phase object [85, 86].

Images collected by TEM, contain both phase and amplitude information [84]. However, the image that is produced is affected by some distortions and aberrations. Factors such as spherical aberrations are caused when the beam is refracted by electromagnetic lenses in different ways such that several focal points are made. Astigmatism is caused by misalignment of electromagnetic and results in blurring. A fast Fourier transform (FFT) of an electron micrograph provides information about the image in the frequency domain. The contrast transfer function (CTF) is a mathematical representation describing some aberrations and the defocus of the TEM image [87]. The FFT of an image is important to monitor during screening because it can be used to identify parameters of the image such as defocus, astigmatism, sample movement (drift), and other factors critical to obtaining the most optimal images. For data collection and image processing, the CTF used to estimate defocus and correct the images for high resolution improvements [88].

1.3.2 Structural methods in cryo-EM

There are four different types of cryo-EM. Each method has its own advantages and disadvantages and is employed to study different types of samples. Electron tomography is used for samples that can often not be averaged such as mitochondria and clathrin cages [89-91]. Electron crystallography makes use of two-dimensional crystals, usually membrane proteins in a lipid bilayer. Similarly, micro-ED is often applied to small 3D micro-crystals [92-94]. Single particle analysis is applied to solubilized protein particles, and is useful for both soluble and membrane proteins where multiple particles exist in the same conformation and can be analyzed via SPA image processing. This method is also applied to samples with some degree of conformational heterogeneity [95]. All of these cryo-EM strategies rely on improving the SNR.

1.3.2.1 Single particle analysis

Recent methods developments (see section 1.3.3) have opened the door to employing SPA on much smaller and heterogeneous protein samples [96, 97]. An electron micrograph suitable for SPA must contain a large number of particles (70-80% of the image) without a contaminating background. Hundreds to thousands of particles are selected and extracted, sometimes using band pass filtering and masking to eliminate any additional background noise [98]. The particles are rotationally and translationally aligned in respect to each other. Multiple iterations of this process results in particles that are optimally aligned for averaging. The resulting averaged image yields a better representation of the particle of interest than any one particle image could give [99]. Due to the random orientations that particles have in solution and usually on the cryo-EM grid, particles are sorted into classes before 3D reconstruction. Using multivariate statistical analysis and hierarchical ascendant classification techniques discussed in the

next section, individual particles are sorted into separate classes [95, 100]. It is theoretically possible to obtain perfect images which could push the size limitations down to 40 kDa [101]. Currently though, samples of approximately 200 kDa require careful, expert analysis to attain a structure at high resolution [102].

1.3.3 Imaging biological samples by TEM

Biological samples are vulnerable to radiation damage by the electron beam because they are mostly made of lighter elements such as hydrogen, carbon, nitrogen, and oxygen [103]. They are also stabilized in an aqueous solution, which would be evaporated in the vacuum environment of the electron microscope. By drying the sample onto the TEM grid before insertion into the column, the specimen can be viewed by TEM, however, the contrast will be very low and the sample will be damaged quickly. Heavy metal salts, such as uranyl acetate or phosphotungstic acid, are used to increase contrast because they scatter the electrons better and create a negative stain projection. One disadvantage to using the negative stain technique is that the resolution is limited [76, 104]. Another problem with negative stains is that the drying of the specimen may cause distortions and flattening [105].

Biological specimens prepared for high resolution data collection by cryo-EM are are vitrified and imaged under low electron doses [106]. This process is done by flash freezing the specimen (usually in liquid ethane) within a thin layer of buffer solution. Flash freezing prevents the water molecules from crystallizing [107, 108]. The contrast produced in cryo-EM is low, and the addition of buffer components such as glycerol, sucrose, and detergents causes further reduction of the overall contrast [109]. However,

several methods are used today to overcome the low signal to noise ratio (SNR). SPA cryo-EM has grown rapidly over the years due to recent advances in high-throughput direct electron detectors and image processing [110-114]. These breakthroughs popularized the method and lead to the recent 2017 Nobel Prize in Chemistry for cryo-EM development.

1.4 Single Particle Analysis of Membrane Proteins

1.4.1 Background

The need to grow crystals is a major bottleneck in both electron and X-ray crystallography. Large macromolecular assemblies can not only be difficult to crystallize, but maintaining activity during crystallization and keeping the protein intact can be another difficult task [76]. Membrane proteins present even more challenges in purification and crystallization methods. Single particle cryo-EM has advantages for these types of problems. In SPA, images of the proteins are taken directly, and computational methods allow for averaging of the particles to achieve a final reconstruction. Therefore, large complexes, highly symmetrical assemblies, and homogenous protein samples are ideal candidates because they can be easily identified and averaged [112]. There are three phases of a single particle cryo-EM project, each one having an essential role in the overall resolution of the final structure (Figure 1.7). The first phase is the biochemistry. It involves the overexpression, purification, and characterization of the protein. Ideally the protein should be as homogenous as possible because the particles will ultimately be averaged. If the sample is highly heterogeneous, due to various conformations or flexible regions, the averaging can result in lower

resolution structures. Often times a membrane protein can lose its activity if the sample undergoes too many purification steps, so finding a balance between high quality data and functional protein will lead to the best strategy for answering relevant questions. In the next phase, the purified sample is screened at a TEM by negative stain. This is a relatively efficient method of determining if the particles are suitable for cryo-EM. Additionally, image processing is performed at this step to obtain a low resolutions model. The final phase involves optimizing vitrification for cryo-EM, collecting large data sets, and processing the data to obtain a structural model which can be analyzed and interpreted to answer fundamental questions. The overall process is non-linear: one or more iterations of optimizing the purification and grid preparation may be required in order for the protein to yield the best possible cryo-EM data.

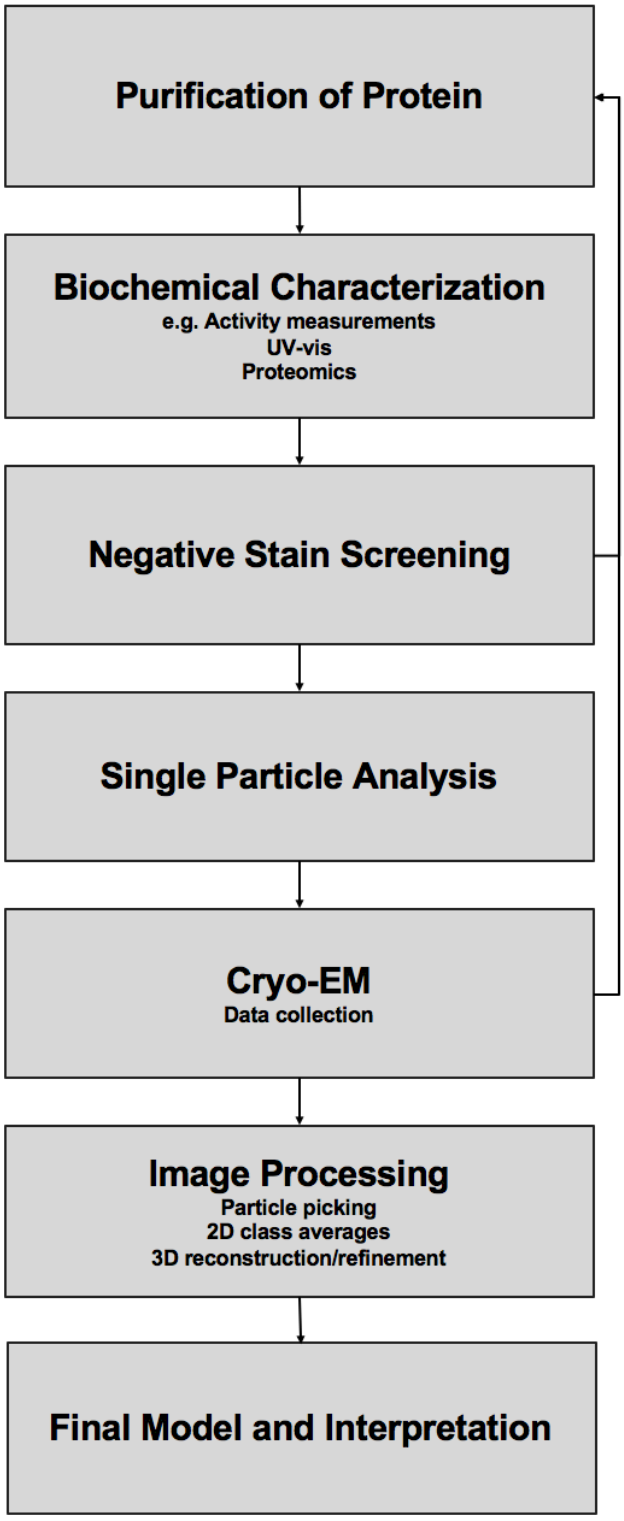


Figure 1.7 Flowchart of single particle cryo-EM.

1.4.2 TEM Screening of single particles by negative stain

Biological molecules absorb and scatter electrons very weakly, leading to low contrast when imaging via TEM. Additionally, the sample cannot withstand the radiation damage for very long [107]. As mentioned previously, freezing the sample in vitreous ice is one method of preserving the sample, but the faster, and more efficient means of screening the purified protein during the optimization phase is by negative stain. This technique is a pivotal step in the overall project.

As described earlier, the contrast in negative stain TEM images is a result of the heavy metal salts that coat the sample. These stains form a protective cast around the biological specimen, preserving the features of the protein so that it may be imaged with adequate contrast. Because the particles are deposited onto a continuous carbon-coated TEM grid (which acts as a flat surface that the protein adheres to), it is not uncommon to see protein with a preferred orientation on the grid.

Often times the concentration for optimal imaging must be determined empirically, especially when working with new samples. Concentrations that are too high may show signs of individual particles with poor separation (Figure 1.8 A). On the other hand, extremely low concentrations will result in sparse, featureless grid squares (Figure 1.8 C). Several trials must be performed in order to find the best concentration for single particle analysis. Though it is not required, particles with uniform structure and very little conformational differences make automated particle picking algorithms feasible. However, resorting to the manual picking is best for more difficult proteins or first time projects.

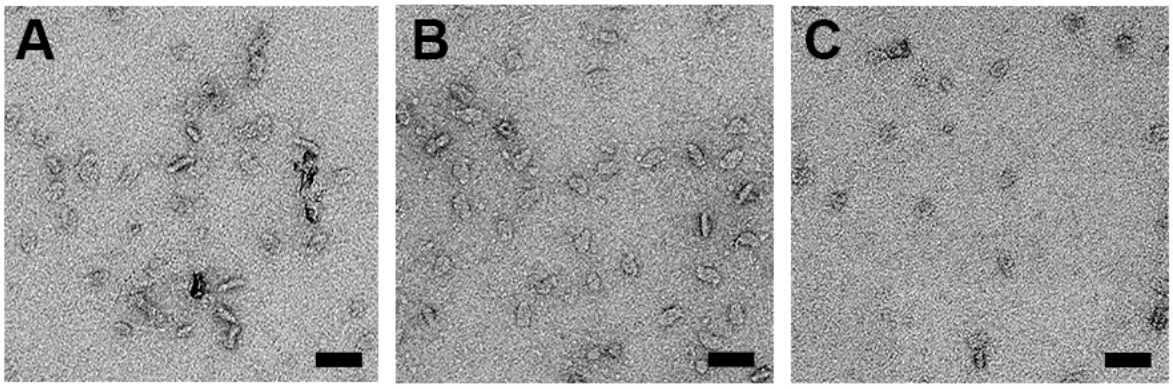


Figure 1.8 Protein concentration is assessed visually by negative stain TEM. Representative TEM images of protein sample on a carbon coated grid. Image (A) illustrates high protein concentration, resulting in difficulties in resolving individual particles. (B) represents an ideal protein concentration suitable for single particle analysis. (C) shows very low particle concentration. Scale bars correspond to 50 nm.

The surface quality of the carbon coated TEM grid is usually hydrophobic, which can affect the abundance of particles that are visualized. A technique called glow discharge, gives the carbon surface of the TEM grid a charged hydrophilic quality when performed under certain conditions. Changing the gases in the atmosphere within the glow discharge unit can create hydrophobic or hydrophilic surfaces, with positive or negative charges [76, 115, 116]. Aside from glow discharging the surface, altering other variables such as the concentration of the sample and/or detergent, temperature, staining times, and pH can all contribute towards achieving good contrast and clearly separated protein particles.

1.4.3 Vitrification of single particles for cryo-EM

After a preliminary structure of the negative stained samples has been determined, cryo-EM preparation is optimized. A metal (usually copper) TEM grid, which has a thin substrate layered on top of the grid squares, such as gold foil or graphene, is used for cryo-EM purposes. This film has an arrangement of holes within each grid square. The goal is to get the protein to be suspended in the holes, leaving just a thin layer of vitreous ice that is not too thick or thin (Figure 1.9). Typically, the sample concentration is much higher in cryo-EM compared to negative stain due to the necessity to get many of the particles to be suspended in the holes of the substrate, however, less than 0.1% of the sample volume is actually used for data collection and this amount is still far less than what is required for crystallography [117].

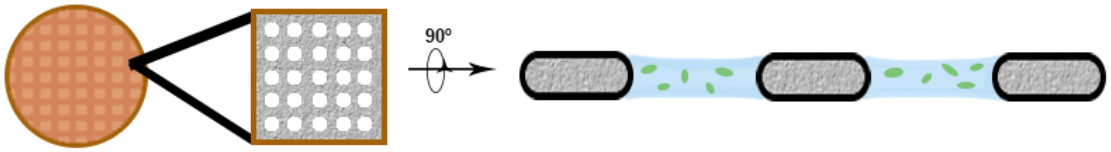


Figure 1.9 cryo-EM grid preparation. A copper mesh grid has a thin layer of carbon on the surface. Within each grid square, there are perforations, either regularly spaced or randomly scattered. Upon vitrification, proteins are frozen in many random orientations within these holes.

There are several ways to prepare a cryo-EM grid. It is possible to blot and plunge using simple home-made devices, but semi-automated methods have been developed by companies [109, 118]. A popular device called a Vitrobot (Thermo Fisher Scientific), used for systematic freezing of grids, allows the regulation of factors such as humidity, ambient temperature, blotting duration, and blot force [119]. Upon a short incubation of the sample on the grid, excess liquid is blotted with filter paper and the grid is quickly plunged into the cryogen. Liquid ethane is typically used as the vitrifying agent due to its high heat capacity, which makes the rate of freezing almost instant and avoids ice crystal formation [119]. The vitrified ice freezes the sample in its native conformation and can even arrest the sample in stages critical to the reaction mechanism [120]. Vitrification preserves positions of atoms within a cage of ice while limiting free radicals and any major changes to the chemical bonds [110]. In addition the particles ideally take on many different orientations when flash frozen. Although, it is sometimes possible for a sample to adopt a preferred orientation within the ice due to some particles having a tendency to be at the air-water interface or clustering near the edges of the holes in the grid [121, 122].

1.4.4 Data collection and image processing

The final step in SPA cryo-EM is the data acquisition and image processing. Depending on the protein size, concentration, and homogeneity, the number of images to be collected can vary. It is typical to collect thousands of micrographs, which yield hundreds of thousands of particles [84, 112]. One reason for the high number of can be sample heterogeneity limiting the number of target particles that are present per image. The primary reason, though, is due to the fact that cryo-EM images yield very low

contrast. The low contrast is a result of the effort to minimize radiation damage (5-20 $e^-/Å^2$). Single particles need to be exposed to a minimum dose coupled with a defocus that allows sufficient contrast to align the particles in an image [84]. The required dose is described in Equation 1.2, where p_{min} is the minimum dose, c is the contrast, D is the particle diameter, and d is the resolution [123].

$$p_{min} \geq \frac{3}{c^2 D d} \qquad \textbf{Equation 1.2}$$

From this equation, it is clear to see that particle size has a role in the critical dosage. In other words, the larger the particle, the easier it is to work with in single particle analysis. However, as improvements in technology are made, smaller and smaller proteins can be studied by cryo-EM.

A technique to minimize electron dose takes advantage of several imaging modes: a low magnification mode, a focusing mode, and an acquisition or expose mode [124]. In low magnification, the electron dosage is very small (often 0.01 $e^-/Å^2$) but allows enough contrast to get an overview of the entire grid to assess qualities such as ice thickness, broken carbon film, large contaminants, etc. In the focus mode, the magnification is increased to the level at which the image would be acquired or higher, however the region of focus is adjacent to the actual area of interest. This allows the focus to be calibrated to obtain sufficient contrast, while sacrificing an area that is not important.

Finally, the image acquisition mode centers over the area of interest at the desired magnification and at a defocus value for suitable contrast [125, 126].

There are several recent advances in technology in the data collection and image processing steps which have helped break resolution barriers in single particle cryo-EM. Changes and advances in detectors have improved the quality of images obtained. Indirect electron detection is usually the process by which transmitted electrons are detected. This requires that primary electrons are converted into photons on a scintillator (such as a CCD chip). Today, direct detection recognizes the primary transmitted electrons via detectors such as ones with a protected, complementary metal-oxide semiconductor (CMOS), which ultimately leads to better resolutions [127-129]. In addition to advances in sensors, movie frames can be collected and used to correct image blurring [128]. Aberrations occur because the illumination of the sample by the electrons causes beam-induced motion. Mechanical drift caused by environmental factors and vibrations also contribute. By collecting images at a high frame rate (10-40 frames per second), it is possible to align the frames together to correct for the slight drift during acquisition [130]. Lastly improvements in image processing software have made the process of 2D averaging and 3D reconstruction user-friendly and accessible for a large range of scientists [131, 132].

Many software packages have been developed for image processing single particles, both open source and proprietary. Some have been developed and optimized for use on a computing cluster, while others have strived for a single workstation interface with CPU or GPU acceleration [131-137]. There are various approaches that each software takes, and typically they share ideologies from each other, however the

underlying concepts are comparable [114]. First, the data set is imported and the CTF is estimated. Then particles are picked, which can be done manually or automatically. Next, the particles are extracted, aligned, and averaged to generate 2D classes. Upon removal of certain unsuitable particle classes, a 3D classification step can be performed followed by refinement to a final reconstruction.

Particle picking can be time consuming, yet is a crucially important experience of the entire image processing pipeline. Getting familiar with the particles of a new sample is necessary in order to assess the overall quality. Manual picking is usually performed as a first step in order to develop a set of 2D class averages which can then be used as a template for automated picking [138]. Automated picking algorithms have been developed in an attempt to make single particle reconstruction a fully automated process. They use cross correlation, variance mapping, intensity comparisons, edge detection, and neural network techniques [139-141]. Using Fourier methods, the cross correlation function (CCF) is calculated between a reference and each micrograph. Peaks in the CCF suggest possible locations of the particles within each micrograph. Despite great efforts in improving these algorithms, sometimes it is still necessary to perform manual picking, especially if the sample is heterogeneous and small. However, it is not absolutely critical if some contaminating particles are selected during a first round of manual or automated picking because unsuitable particles can be filtered out in later steps. Ultimately one has to decide whether to perform a somewhat biased, yet more accurate approach to picking (manual) or an unbiased, yet potentially contaminating method (automated). This decision is best made on a project by project basis.

Particles must be classified into various groups based on factors such as orientation, size, and conformation. Using multivariate data analysis, it is possible to represent the overall images as vector points in order to reduce the data into a low-dimensional representation. Together, these points provide a "cloud" of points in factor space which will cluster according to certain trends, such as conformational shape, ligand bound/unbound, open/closed states, etc. [100]. Then, using the relative closeness of these points in factor space, a classification technique is used, such as hierarchical ascendant classification or K-means clustering [84, 142].

Upon classification, the ideal classes representing different orientations of the protein are selected, while other classes are deemed extraneous. These classes represent all the 2D particle images that will be used for the 3D reconstruction. Due to irregularities in the particle images, real-space reconstruction techniques can be used, though Fourier space reconstruction methods have some advantages [131, 137, 138, 142]. Weighted, or filtered back-projection is a common reconstruction method which employs a Fourier component. First, 2D projections of the protein are obtained at various angles. If one back-projects these images to create the original image, one obtains a "smear" of what the original object was. This is due to the low frequency data in Fourier space is sampled much more densely than the higher frequencies, which contain finer details of the object. By taking a Fourier transform of each projection, applying a weighting factor or filter to the spectrum to increase high frequency data, and then applying an inverse Fourier transform, the result is a series of projections with both high and low frequency information which delivers sharp and smooth details from the original object. Then, upon

recombining and back-projecting the images, one obtains a reconstructed object that is more similar to the original [143].

Real-space reconstruction techniques are iterative, so they are relatively slow in terms of speed of processing. Fourier algorithms offer a one to two orders of magnitude increase in speed over typical back-projection methods. They attempt to find the values of the 3D Fourier transform on a 3D Cartesian grid. Based on the Central Slice Theorem, or projection theorem, one can combine the 2D Fourier transforms of many experimental projections into the Cartesian 3D Fourier transform. Given a reference model, usually obtained from another similar or previously solved structure, one can use this to computationally reference the data into the correct locations and orientations.

After an initial 3D reconstruction of the protein is completed, further refinements are necessary due to resolution loss from creating the initial model. Angular refinements must be made, typically by using a 3D reference model. Reference projections are created from the reference model and the calculated projections are matched with actual experimental projections until a high CCF score is achieved. Each iteration of matching is used to compute a new density map based on the experimental data [84].

When a resolution of a structure is reported as a number, that unit is described as the maximum spatial frequency at which information can be considered reliable [144]. Resolution in X-ray crystallography is calculated based on diffraction data [145]. In single particle cryo-EM, one can get a rough first, idea of resolution based on the Thon rings in a Fourier transform of a micrograph after frame alignment [146]. In order to properly estimate the consistency of a structure solved by single particle cryo-EM

methods, it is common to use Fourier Shell Correlation (FSC) to measure the internal consistency of the data [147-149]. Two 3D volumes can be converted in Fourier space and their normalized cross-correlation coefficient can be plotted as a function of spatial frequency. The gold standard method is to split the experimental data by two equal halves and estimate the resolution of each until they converge to a certain point.

1.5 Significance

Photosystem II evolving in cyanobacteria lead to the terraforming of Earth which radically changed the environment towards supporting aerobic life. The abundance of water as the electron source for these photosynthetic species aided in the rapid growth and colonization, which eventually filled the atmosphere with oxygen and created an ozone layer to protect life from harsh UV-radiation [58]. This entire process could not be possible without PSII, and it is therefore important to analyze the structure and function in detail. The significance of studying PSII has implications in solar energy research, agriculture, and alternative energy in the form of biomimetics. Sunlight indirectly or directly powers most processes on Earth and photosynthesis is the starting point of the conversion of this energy. Found in all photosynthetic life forms, PSII is the macromolecular complex that drives the conversion of water into protons and oxygen gas. It is the reason why we have fossil fuels today and continues to sustain life on this planet [150].

Biochemical and biophysical research on PSII has provided key insights into the blueprint of how nature evolved a way to utilize light energy. Structural information of a highly active PSII is necessary to fully understand the detailed reaction mechanism. PSII

has proven to be an ideal candidate for cryo-EM and it is critical to study the highly active complex. Structural heterogeneity in a protein sample exists either because of conformational differences, flexibility of proteins, or differences in complex sizes (i.e. presence or absence of specific subunits). Single particle cryo-EM is the only structural techniques that permits some degree of heterogeneity [151]. Though the ease of SPA is greatly improved with a mono-disperse sample, the work of this thesis shows that heterogeneous samples of highly active PSII can be studied by cryo-EM. Membrane proteins often have purification limitations, whether it be a lack of abundance or lack of homogeneity. However, these protein structures are important for drug design and biotechnology. Thus, these results will contribute to and facilitate cryo-EM studies of this important group of proteins.

CHAPTER 2. BIOCHEMICAL CHARACTERIZATION OF PHOTOSYSTEM II

2.1 Introduction

A key step, and a very frequently the largest hurdle to single particle cryo-EM, is the biochemical isolation suitable for successful data collection. In addition, biochemical characterization and retention of function is critical to all structural approaches. Being a large, macromolecular complex, plant PSII has many challenges in its purification process. Retention of activity was the utmost important factor when purifying the protein and involved measuring the activity before negative stain TEM and cryo-EM experiments. Subunit composition was determined by mass spectrometry and SDS-PAGE, and spectroscopic methods were used to assess the presence of antenna proteins. In this chapter, the discussion of the different purification methods that yielded various types of PSII complexes will provide an understanding of the major differences between current structures of PSII and the sample we produce.

A robust and popular method of purification of plant PSII has been published, which yields PSII rich thylakoid membrane fragments, called the BBY PSII, though sometimes this method refers to the sample as BBY particles. These PSII-enriched membranes are solubilized using Triton X-100 (Sigma. T9284-500ML. Lot# 053K00262V) detergent [38, 57]. This method is then moved to an octyl β-D-thioglucopyranoside (OTG) core preparation yielding PSII dimers without the light harvesting antenna proteins. This step is called the OTG-PSII [152]. This method of producing highly active PSII core complexes was seen in previous studies and shown to

evolve oxygen greater than 1000 μmol O_2 mg chl^{-1} hr^{-1} [153]. Other purification methods result in oxygen evolving rates below 300 μmol O_2 mg chl^{-1} hr^{-1} (Chapter 1, Table 1.3).

In order to achieve the suitable conditions for TEM imaging of the OTG-PSII cores, it was found that a buffer exchange to dodecyl β-D-maltoside (β-DDM) detergent was necessary. The OTG-PSII did not produce sufficient contrast to image the particles and had a predominant orientation on the TEM grid. Once the sample was exchanged into β-DDM buffer, hereafter called the OPDM sample, the contrast improved and the imaging was reproducible. As a control for the active PSII sample, we also purified PSII-LHCII supercomplexes by sucrose gradient ultracentrifugation, which separates the mixture of complexes in the BBY step based on size [153, 154]. However, it was shown that the activity of these samples did not maintain high rates compared to the OPDM preparation.

The work presented in this chapter shows the biochemical studies of a purification method that is optimized for TEM and maintains high oxygen evolution rates. The sacrifice for such high rates comes at the cost of complex heterogeneity, which will be discussed in the next chapter. The results described here outline a strategy for a robust, reproducible method of producing stable, highly active PSII for cryo-EM analysis. In addition, it provides the foundation for a large range of additional structure and function experiments on the PSII complex.

2.2 Results

2.2.1 Purification and isolation of PSII using two methods

In both the inactive and active purification preparations, spinach (*Spinacea oleracea*) leaves were blended, filtered, and subjected to several low speed centrifugation steps in order to remove larger cell debris. It was found to be particularly important to add Triton X-100 slowly during the incubation step to avoid over-solubilization while mixing. Upon the isolation of PSII-rich BBY membranes, oxygen assays were measured to ensure the activity is still retained (Table 2.1).

The inactive C_2S_2 complexes were made by solubilizing the BBY membranes with β-DDM and layering the solution carefully onto a sucrose gradient. A 20% solution of sucrose was found to be sufficient in generating a gradient which could separate the proteins well enough to be extracted. Upon high speed centrifugation the various components of the BBY sample were separated throughout the gradient (Figure 2.1). The bands were labeled B1-B11. According to previous literature and some spectroscopic and TEM analyses, the bands were identified as the following. B1 contained the free pigments. B2 contained monomeric light harvesting complex. The dark green band, B3, had the LHCII trimers. B4 contained the mixtures of the trimers and the CP24, CP26, and CP29 antenna proteins. B5 contained monomeric PSII, and B6 contained monomeric PSII with a single LHCII trimer. Bands 7-9 were more difficult to isolate, however they represent the various combinations of the C_2 dimer with one or two copies of the strongly bound LHCII trimer and the CP26 and CP29 proteins. Finally, B10 and B11 contained the $C_2S_2M_2$ megacomplex, with the asymmetrically or symmetrically bound M-LHCII

trimer and CP24. The oxygen evolving activity of the sample after the sucrose gradient step was low (Table 2.1). A representative overview of band 9 extract depicts the C_2S_2-type PSII complexes in TEM (Figure 2.2A).

To make the highly active OTG-PSII cores and OPDM sample, the BBY membranes were further solubilized in OTG detergent to separate the PSII core dimers from the thylakoid membranes. Again, it was important to slowly add the OTG to allow for gradual solubilization. Treatment with $MgCl_2$ was necessary for the removal of the LHCII, CP26 and CP29 proteins, though some residual antenna proteins still remained in the sample based of spectroscopic measurements (Figure 2.6). However, when imaging by TEM (Figure 2.2B), it was clear that the majority of the complexes were of the C_2 type and these could selectively be chosen during image processing, which will be discussed in Chapter 3. To produce OPDM cores suitable for TEM, a β-DDM buffer exchange proved to be necessary. This buffer exchange yielded C_2 type PSII particles with high contrast in negative stain that could not be visualized at equally high contrast when imaging the OTG-PSII cores. These results are described in Chapter 3. Any modifications made to the samples, including the dilutions for negative stain and cryo-EM, were tested for oxygen evolution rates to ensure high activity was still maintained.

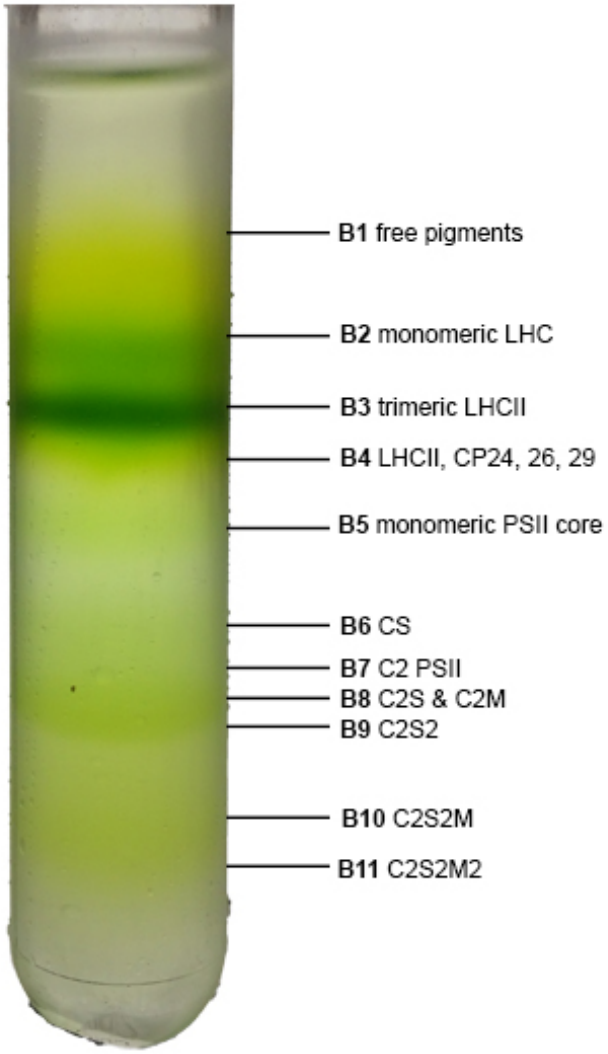


Figure 2.1 Isolation of PSII by sucrose gradient ultracentrifugation. Bands 1-11 are labeled and their components are listed. Nomenclature for the types of PSII complexes are described in Figure 1.4.

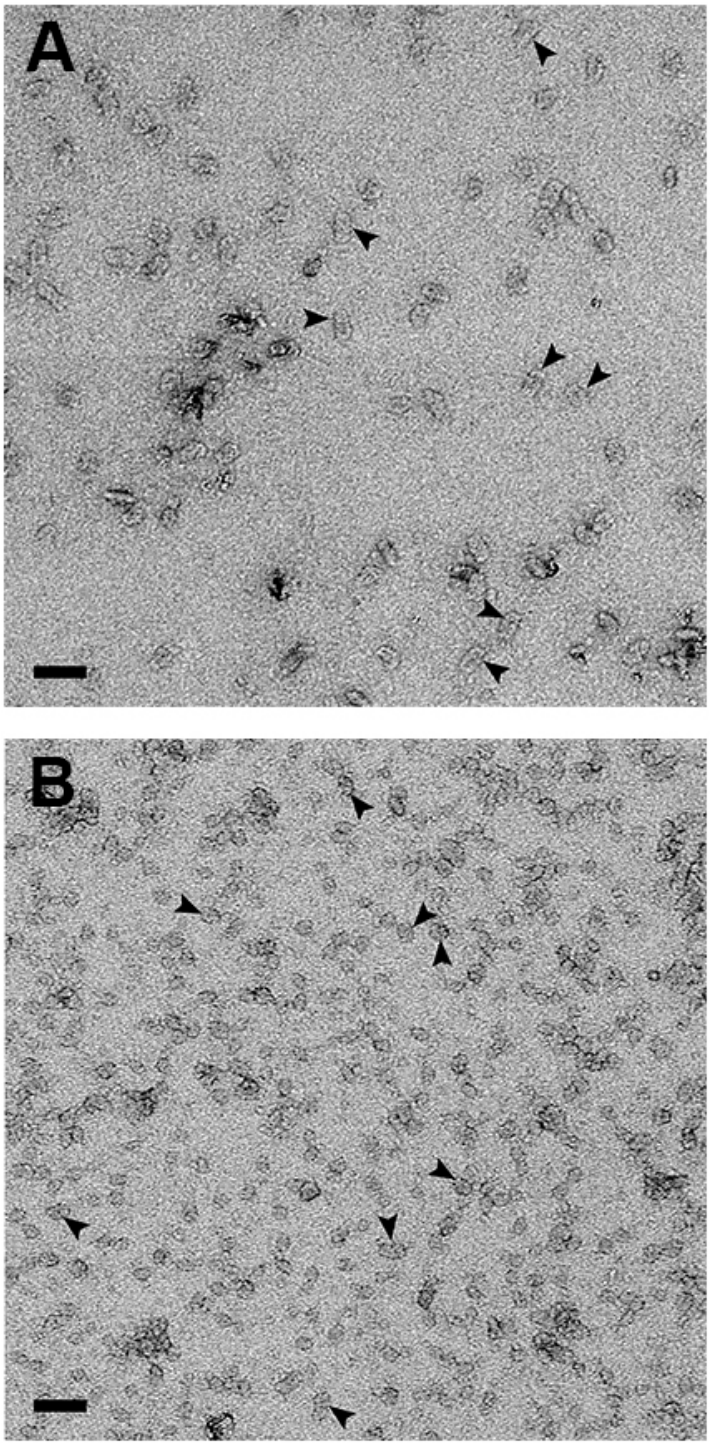


Figure 2.2 Representative micrographs of negatively stained PSII particles. (A) The sucrose gradient purification produces C_2S_2 complexes (top views indicated by black arrow heads). (B) OPDM single particles show largely C_2-type PSII (top views indicated by black arrow heads). Scale bar corresponds to 50 nm.

2.2.2 Oxygen evolution activity of PSII in various purifications

At three key points in the purification, sample was extracted and tested for oxygen evolution (Table 2.1). At the BBY step, our activity rates tended to average approximately 800 μmol O_2 mg chl^{-1} hr^{-1}. The standard rates reported for BBY PSII preparations at pH 6 are 300 μmol O_2 mg chl^{-1} hr^{-1} [57]. At the OTG-PSII step, the rates were 1200-1300 μmol O_2 mg chl^{-1} hr^{-1}. Finally, the OPDM modification was shown to have little to no effect on the rates compared to the OTG-PSII cores. The sucrose gradient PSII had rates similar to the C_2S_2 supercomplexes from the 2016 structure [18]. The addition of $CaCl_2$ is a means to test if a decrease in activity is due to the loss of extrinsic subunits (explained further in the Discussion section). $CaCl_2$ added to both the OTG-PSII and the OPDM showed no significant difference (Table 2.1).

In order to normalize the rates to a standard unit for comparison across all types of PSII, it was necessary to convert the rates to μmol O_2 evolved per mol reaction center per hour. The rates for the BBY, OTG-PSII, OPDM, and the samples from the recent plant PSII structures are summarized in Table 2.3.

Table 2.1 Oxygen evolution rates observed in PSII at various steps in the purification. All buffer conditions kept at pH 6.0 except for sucrose gradient PSII at pH 7.5.

Purification Method	O_2 Evolving Activity (μmol O_2 mg chl^{-1} hr^{-1})
BBY	1000 ± 100
OTG-PSII	1300 ± 100
OTG-PSII (w/$CaCl_2$)	1300 ± 50
OPDM	1200 ± 200
OPDM (w/$CaCl_2$)	1100 ± 200
sucrose gradient PSII	100 ± 20

The OTG-PSII cores did not have suitably high contrast in negative stain (Chapter 3), so it was necessary to alter the buffer conditions to optimize the sample for negative stain, and ultimately cryo-EM. Table 2.2 describes the oxygen evolution rates of various conditions. The control OTG-PSII cores evolved at 1460 μmol O_2 mg chl^{-1} hr^{-1} and all other conditions were compared to this rate. The pH 7.5 buffer exchange indicated a loss of extrinsic subunits because the rates dropped to 570 μmol O_2 mg chl^{-1} hr^{-1}. The addition

of $CaCl_2$ rescued the activity to 1320 μmol O_2 mg chl^{-1} hr^{-1}. The addition of 0.008% β-DDM did not have a significant affect on the oxygen activity, but when combining pH 7.5 and β-DDM, the rates dropped, indicating that pH has a large impact on the retention of extrinsic subunits and activity. Finally, 0.6% β-DDM buffer showed no significant decrease in oxygen evolution rates.

Table 2.2 Oxygen evolution rates in OTG-PSII cores with different buffer treatments.

Buffer conditions	O_2 Evolving Activity (μmol O_2 ·mg chl^{-1} · hr^{-1})
OTG-PSII	1460
pH 7.5	570
pH 7.5 (w/$CaCl_2$)	1190
0.008% DDM	1320
0.008% DDM+pH 7.5	600
0.008% DDM+pH 7.5 (w/$CaCl_2$)	980
0.6% DDM	1280

Table 2.3 Normalization of oxygen activity rates in various PSII preparations.

PSII type	Year	Plant Species	O_2 Evolving Activity ($\mu mol\ O_2 \cdot mg\ chl^{-1} \cdot hr^{-1}$)	Normalized Activity ($\mu mol\ O_2 \cdot mol\ rxn\ center^{-1} \cdot hr^{-1}$)
C_2, OPDM	2018	*Spinacia oleracea*	1200	5.9×10^{10}
C_2, OTG-PSII	2018	*Spinacia oleracea*	1300	6.4×10^{10}
Thylakoid membranes, BBY	2018	*Spinacia oleracea*	800	20×10^{10}
C_2S_2	2016	*Spinacia oleracea*	75	0.7×10^{10}
$C_2S_2M_2$ (stacked)	2017	*Pisum sativum*	288	4.0×10^{10}
$C_2S_2M_2$ (unstacked)	2017	*Pisum sativum*	117	1.6×10^{10}
$C_2S_2M_2$	2017	*Arabidopsis thaliana*	not reported	N/A
$C_2S_2M_2$ (stacked)	2017	*Pisum sativum*	not reported	N/A

2.2.3 Subunit characterization by biochemical techniques

In order to understand the composition of the samples that were purified, biochemical assays TEM were used to verify that we had the components necessary for an active PSII complex. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was an important tool to account for the presence of critical subunits in the PSII complex. In addition, PSII is rich in chlorophyll molecules, which have unique signatures in the red and blue regions, making spectroscopic assays a useful technique in determining if antenna proteins are present in the complex. Finally, analysis by mass spectrometry verified the subunits identified by SDS-PAGE and elucidated smaller subunits that went undetected by the Coomassie staining.

2.2.3.1 SDS-PAGE

A summary of the major subunits detected by Neville SDS-PAGE analysis of PSII is shown in Table 2.4. The gel images of the three different types of PSII in our purification is shown in Figure 2.3. The major core intrinsic proteins (CP43, CP47, D1, D2) were identified in all three active PSII types, as well as the extrinsic polypeptides (PsbO, PsbP, and PsbQ). The additional bands present in the gel are from the smaller intrinsic proteins and the antenna proteins (LHCII, CP26, CP29, CP24). Figure 2.4 describes the SDS-PAGE results from varying the detergent concentration in OPDM samples. The control sample was OTG-PSII cores. The 0.008% β-DDM OPDM sample was tested to simulate the approximate amount of β-DDM used in the sucrose gradient purification. A 0.6% β-DDM OPDM sample was also tested. These results show that the

amount of β-DDM did not seem to affect the subunit composition. The final detergent concentration used for cryo-EM data collection was 0.3% β-DDM (Chapter 3).

Table 2.4 Summary of major plant PSII subunits identifiable by SDS-PAGE.

Subunit	Intrinsic/Extrinsic	Molecular Weight (kDa) in SDS-PAGE
D1	Intrinsic	30
D2	Intrinsic	30
CP43	Intrinsic	43
CP47	Intrinsic	47
PsbO	Extrinsic	33
PsbP	Extrinsic	23/24
PsbQ	Extrinsic	17/18

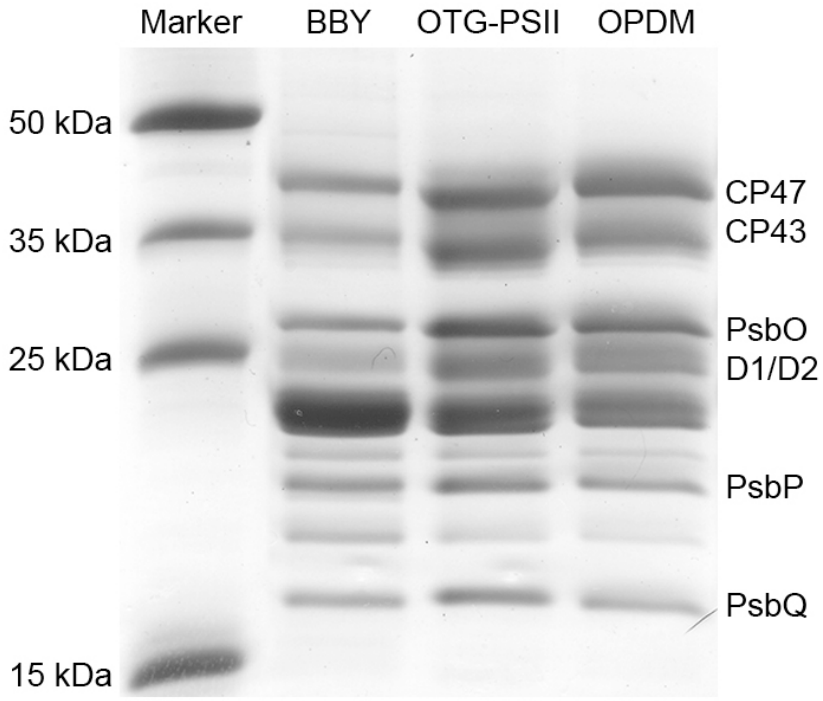


Figure 2.3 Urea SDS-PAGE analysis of three stages in the highly active PSII purification. The first lane represents PSII membranes (BBY) isolated by Triton X-100. The second lane is PSII isolated by OTG. The third lane is PSII re-solubilized in DDM. All samples contain the core intrinsic proteins (CP47, CP43, D1, D2) and extrinsic subunits (PsbO, PsbP, PsbQ). Other non-labeled bands are antenna proteins and smaller intrinsic polypeptides.

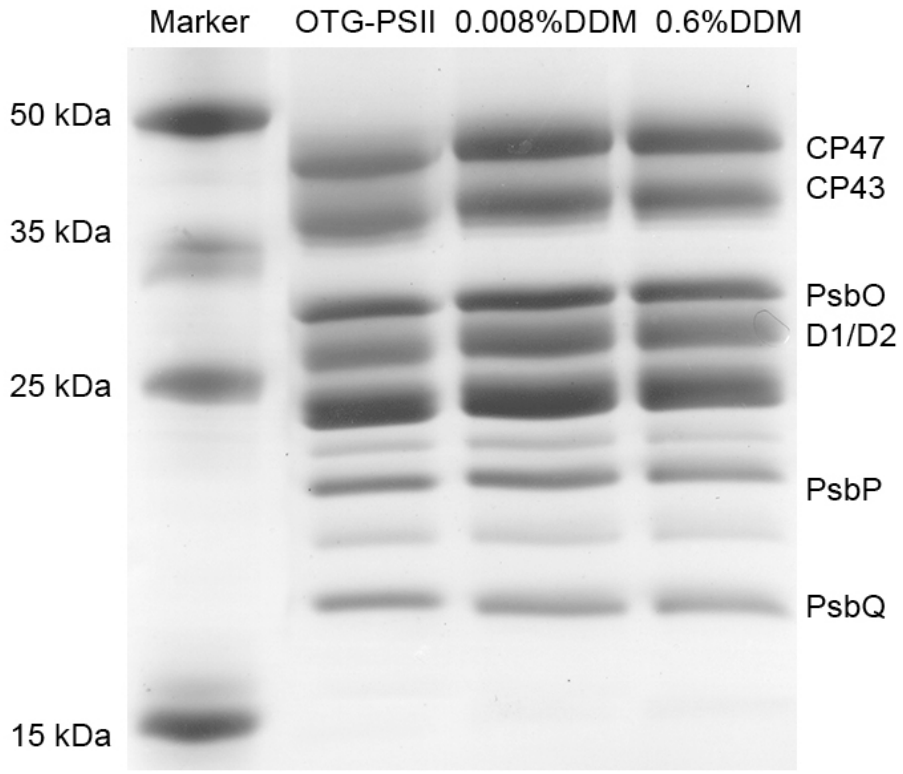


Figure 2.4 Urea SDS-PAGE analysis of OPDM in varying DDM concentrations. The first lane represents the OTG-PSII as a control. The second lane is OPDM solubilized with 0.008% DDM, and the third lane is OPDM in 0.6% DDM.

2.2.3.2 UV-vis absorption spectroscopy

The protein extracted from the sucrose gradient was analyzed by ultraviolet-visible (UV-vis) absorption spectroscopy (Figure 2.5). The spectra were normalized to a peak at approximately 663 nm, which is one of the peaks where chlorophyll *a* absorbs the most. As described in Figure 2.5, the B3 and B5 bands contain trimeric LHCII and monomeric PSII core proteins, respectively. Chlorophyll *b* exists in the antenna proteins, and it absorbs at around 640 nm, giving a characteristic "shoulder" to the spectra for bands B3 and B5. Monomeric PSII does not contain any chlorophyll *b*, but because of its

closeness in size to CS complexes (monomeric PSII + LHCII trimer), it is likely that this sample also contains a mixture of core and antenna proteins, explaining the presence of the chlorophyll *b*. In addition to the 640 nm region, chlorophyll *b* presence can also be detected near the 460 nm region. The lower bands in the gradient, (B7, B8, B9) contain much less chlorophyll *b*. B7 is made of C_2 complexes, B8 is made of C_2S and C_2M, and B9 is made of C_2S_2. Again, the presence of a higher peak at around 460 nm (chlorophyll *b*) for B9 is seen.

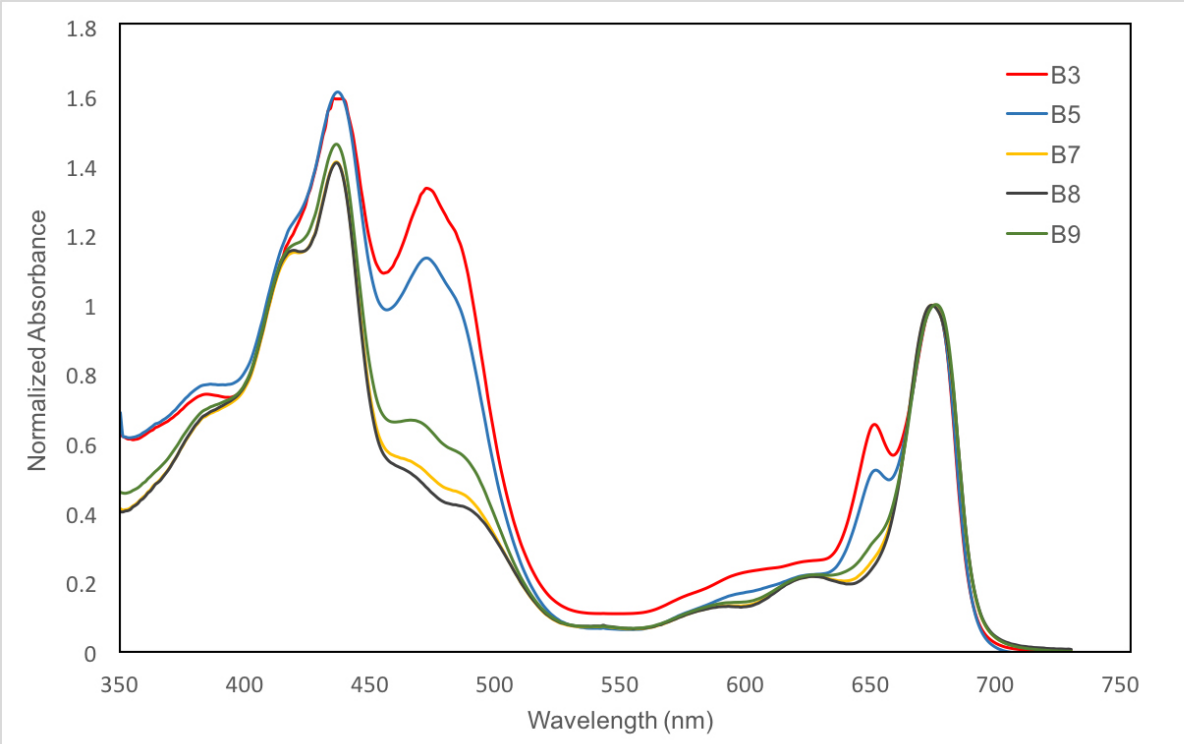


Figure 2.5 Normalized UV-vis absorption spectra of several sucrose gradient fractions. Bands 3, 5, 7, 8, and 9 were extracted from a sucrose gradient preparation and analyzed. The spectra were baseline corrected and normalized to the maximum in the red region (~660 nm).

The BBY, OTG-PSII, and OPDM active PSII samples were also measured by UV-vis absorption spectroscopy. These samples were not subjected to a sucrose gradient. Therefore, they were more heterogeneous compared to the sucrose gradient PSII, however, the global abundance of chlorophyll types was easily identified. In Figure 2.6, the BBY contains the shoulder at around 650 nm and a peak at 460 nm. The OTG-PSII cores and the OPDM sample do not have as pronounced of a peak in this region. The spectra were normalized to the 663nm region.

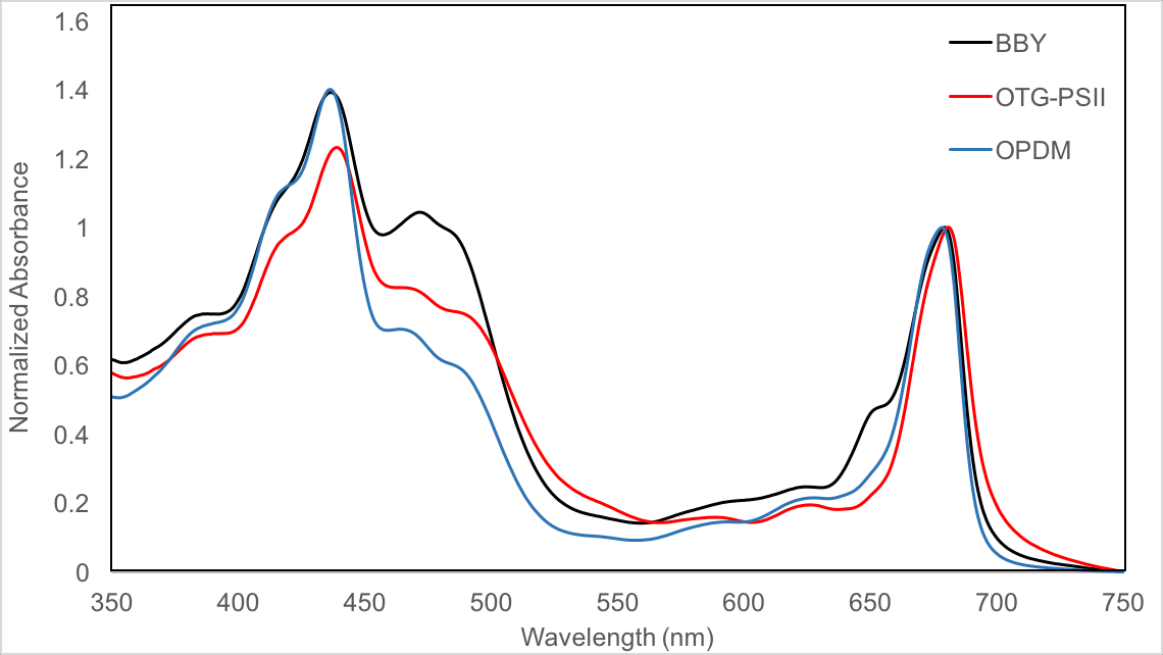


Figure 2.6 Normalized UV-vis absorption spectra of three PSII purification steps. Three points in our purification of C_2 PSII were analyzed for chlorophyll content. The BBY, OTG-PSII, and OPDM samples all had high oxygen evolution rates. The spectra were baseline corrected and normalized to the maximum in the red region (~660 nm).

2.2.3.3 Mass spectrometry

Mass spectrometry was performed in order to assess the smaller subunits that could not be identified with Coomassie stained SDS-PAGE. In addition, the results confirmed the presence of the major intrinsic and extrinsic subunits. The OPDM sample was trypsin digested and run via tandem mass spectrometry. The results of the peptides identified are presented in Table 2.5. The trypsin typically cuts the protein at lysine or arginine residues, indicated by the start and end amino acids in brackets. Modifications typically existed as oxidations or carbamidomethylations. Peptide spectral matches (PSM) are the total number of identified spectra that correspond to the peptides of a specific protein. The PSM value may be higher than the number of peptides identified for high-scoring proteins because peptides may be identified repeatedly. Spectral counting is based on the assumption that the sampling rate of a peptide is related to the abundance of the peptide represented by its precursor ion in the sample mixture [155, 156]. Because larger proteins are expected to produce more peptides, one can obtain a normalized spectral abundance factor (NSAF), which involves normalizing for the protein length and total number of MS/MS spectra. The greater the number of PSMs per unit length, the more abundant that subunit is in the PSII sample.

Table 2.5 Tandem mass spectrometry data identifying peptides of subunits specific to PSII.

Peptide	Subunit	PSM	Modification	Relative abundance (PSM/unit length)
[R].FGQEEETYNIVAAHGYFGR.[L]	D1, PsbA	10		1.69
[R].VINTWADIINR.[A]	D1, PsbA	14		
[R].ETTENESANEGYR.[F]	D1, PsbA	66		
[R].ANLGMEVmHER.[N]	D1, PsbA	217	1xOxidation [M]	
[R].ANLGMEVMHER.[N]	D1, PsbA	62		
[R].FcNWITSTENR.[L]	D1, PsbA	12	1xCarbamidomethyl [C2]	
[R].ESESLWGR.[F]	D1, PsbA	22		
[R].LIFQYASFNNSR.[S]	D1, PsbA	6		
[R].ANLGmEVmHER.[N]	D1, PsbA	143	2xOxidation [M5; M8]	
[R].RESESLWGR.[F]	D1, PsbA	21		
[R].EWELSFR.[L]	D1, PsbA	7		
[R].VSAGLAENQSFSEAWSK.[I]	CP47, PsbB	33		0.45
[R].RVSAGLAENQSFSEAWSK.[I]	CP47, PsbB	26		
[R].MPTFFETFPVVLIDGDGIVR.[A]	CP47, PsbB	4		
[R].YQWDQGYFQQEIYR.[R]	CP47, PsbB	10		
[R].AGSMDNGDGIAVGWLGHPIFR.[D]	CP47, PsbB	3		
[R].AQLGEIFELDRATLK.[S]	CP47, PsbB	1		
[R].RAQLGEIFELDR.[A]	CP47, PsbB	8		
[K].IPEKLAFYDYIGNNPAK.[G]	CP47, PsbB	4		
[R].VHTVVLNDPGR.[L]	CP47, PsbB	65		
[R].AQLGEIFELDR.[A]	CP47, PsbB	12		
[K].LAFYDYIGNNPAK.[G]	CP47, PsbB	5		

Table 2.5 continued.

Peptide	Subunit	PSM	Modification	Relative abundance (PSM/unit length)
[R].VSAGLAENQSFSEAWSKIPEK.[L]	CP47, PsbB	2		0.45
[R].TGKPSLDLPK.[I]	CP47, PsbB	18		
[R].RMPTFFETFPVVLIDGDGIVR.[A]	CP47, PsbB	3		
[R].QGMFVIPFMTR.[L]	CP47, PsbB	3		
[R].RMPTFFETFPVVLIDGDGIVR.[A]	CP47, PsbB	1	1xOxidation [M2]	
[R].QGMFVIPFMTR.[L]	CP47, PsbB	1	1xOxidation [M9]	
[K].IGDPTTR.[R]	CP47, PsbB	21		
[R].ATLKSDGVFR.[S]	CP47, PsbB	8		
[R].QGMFVIPFMTR.[L]	CP47, PsbB	1	2xOxidation [M3; M9]	
[R].YQWDQGYFQQEIYRR.[V]	CP47, PsbB	2		
[R].SAEYMTHAPLGSLNSVGGVATEINAVNYVSPR.[S]	CP43, PsbC	1	1xOxidation [M5]	0.72
[R].SAEYMTHAPLGSLNSVGGVATEINAVNYVSPR.[S]	CP43, PsbC	2		
[R].DQETTGFAWWAGNAR.[L]	CP43, PsbC	6		
[R].LGANVGSAQGPTGLGK.[Y]	CP43, PsbC	48		
[R].SPTGEVIFGGETMR.[F]	CP43, PsbC	24		
[K].GIDRDFEPVLSMTPLN.[-]	CP43, PsbC	12		
[R].GPNGLDLSR.[L]	CP43, PsbC	13		
[R].SPTGEVIFGGETMR.[F]	CP43, PsbC	54	1xOxidation [M13]	
[R].AAAAGFEK.[G]	CP43, PsbC	69		
[K].KDIQPWQER.[R]	CP43, PsbC	51		
[R].APWLEPLR.[G]	CP43, PsbC	8		

Table 2.5 continued.

Peptide	Subunit	PSM	Modification	Relative abundance (PSM/unit length)
[K].DIQPWQER.[R]	CP43, PsbC	27		0.72
[R].DFEPVLSMTPLN.[-]	CP43, PsbC	3		
[K].GIDRDFEPVLSMTPLN.[-]	CP43, PsbC	7	1xOxidation [M12]	
[R].DFEPVLSMTPLN.[-]	CP43, PsbC	1	1xOxidation [M8]	
[K].DIQPWQERR.[S]	CP43, PsbC	13		
[R].APWLEPLRGPNGLDLSR.[L]	CP43, PsbC	3		
[R].AFNPTQAEETYSMVTANR.[F]	D2, PsbD	12		0.41
[R].AFNPTQAEETYSMVTANR.[F]	D2, PsbD	17	1xOxidation [M13]	
[R].AWMAAQDQPHENLIFPEEVLPR.[G]	D2, PsbD	9		
[R].AWMAAQDQPHENLIFPEEVLPR.[G]	D2, PsbD	4	1xOxidation [M3]	
[R].AAEDPEFETFYTK.[N]	D2, PsbD	20		
[R].AYDFVSQEIR.[A]	D2, PsbD	19		
[K].DLFDSMDDWLR.[R]	D2, PsbD	2		
[K].DEKDLFDSMDDWLR.[R]	D2, PsbD	1		
[K].NILLNEGIR.[A]	D2, PsbD	24		
[R].FDSLEQLDEFSR.[S]	cytochrome b559 PsbE	9		0.46
[R].QGIPLITGR.[F]	cytochrome b559 PsbE	25		
[R].SFADIITSIR.[Y]	cytochrome b559 PsbE	4		
[R].TYPIFTVR.[W]	cytochrome b559 PsbF	8		0.21
[R].SRPKPTTVGALLKPLNSEYGK.[V]	PsbH	37		0.55
[M].ATQTVESSSR.[S]	PsbH	3		
[-].MTQSNPNEQNVELNR.[T]	PsbL	5		0.97
[-].MTQSNPNEQNVELNR.[T]	PsbL	6		

Table 2.5 continued.

Peptide	Subunit	PSM	Modification	Relative abundance (PSM/unit length)
[M].TQSNPNEQNVELNR.[T]	PsbL	26		0.97
[K].FEEKDGIDYAAVTVQLPGGER.[V]	PsbO	14		3.64
[R].LTYTLDEIEGPFEVSSDGTVK.[F]	PsbO	27		
[K].GTGTANQCPTVEGGVDSFAFKPGK.[Y]	PsbO	59	1xCarbamidomethyl [C8]	
[K].DGIDYAAVTVQLPGGER.[V]	PsbO	19		
[K].GRGGSTGYDNAVALPAGGR.[G]	PsbO	9		
[R].GGSTGYDNAVALPAGGR.[G]	PsbO	73		
[K].QLVASGKPESFSGDFLVPSYR.[G]	PsbO	24		
[K].NSGPDFQNTK.[L]	PsbO	153		
[K].AEGISKNSGPDFQNTK.[L]	PsbO	27		
[R].GDEEELQKENNK.[N]	PsbO	163		
[K].NSGPDFQNTKLMTR.[L]	PsbO	3		
[K].NSGPDFQNTKLMTR.[L]	PsbO	6	1xOxidation [M12]	
[R].GGSTGYDNAVALPAGGRGDEEELQK.[E]	PsbO	17		
[K].GTGTANQCPTVEGGVDSFAFKPGK.[Y]	PsbO	1		
[K].TYLEVKGTGTANQCPTVEGGVDSFAFKPGK.[Y]	PsbO	2	1xCarbamidomethyl [C14]	
[K].RLTYDEIQSK.[T]	PsbO	104		
[K].GTITLSVTSSKPETGEVIGVFQSLQPSDTDLGAK.[V]	PsbO	12		
[R].LTYDEIQSK.[T]	PsbO	109		
[K].KFCLEPTK.[F]	PsbO	72	1xCarbamidomethyl [C3]	
[K].GTGTANQCPTVEGGVDSFAFKPGKYTAK.[K]	PsbO	5	1xCarbamidomethyl [C8]	
[R].GDEEELQK.[E]	PsbO	131		
[R].GSSFLDPK.[G]	PsbO	105		

Table 2.5 continued.

Peptide	Subunit	PSM	Modification	Relative abundance (PSM/unit length)
[K].FCLEPTK.[F]	PsbO	56	1xCarbamidomethyl [C2]	3.64
[R].VPFLFTIK.[Q]	PsbO	13		
[R].LTYDEIQSKTYLEVK.[G]	PsbO	1		
[K].FAVKAEGISK.[N]	PsbO	1		
[K].FCLEPTKFAVK.[A]	PsbO	2	1xCarbamidomethyl [C2]	
[K].TDSEGGFDSGVVASANVLESSTPVVDGK.[Q]	PsbP	5		1.58
[R].YEDNFDATSNLSVLVQPTDKK.[S]	PsbP	14		
[R].YEDNFDATSNLSVLVQPTDK.[K]	PsbP	18		
[K].SITDFGSPEDFLSQVDYLLGK.[Q]	PsbP	11		
[K].KNTEFMPYNGDGFK.[L]	PsbP	13	1xOxidation [M6]	
[K].KNTEFMPYNGDGFK.[L]	PsbP	14		
[R].TADGDEGGKHQVIAATVK.[D]	PsbP	42		
[K].HQVIAATVK.[D]	PsbP	137		
[K].KFVESATSSFSVA.[-]	PsbP	35		
[K].NTEFMPYNGDGFK.[L]	PsbP	8	1xOxidation [M5]	
[K].QYYSITVLTR.[T]	PsbP	8		
[K].EKEFPGQVLR.[Y]	PsbP	44		
[K].LLVPSKWNPSK.[E]	PsbP	8		
[K].EFPGQVLR.[Y]	PsbP	49		
[K].NTEFMPYNGDGFK.[L]	PsbP	7		
[K].FVESATSSFSVA.[-]	PsbP	7		

Table 2.5 continued

Peptide	Subunit	PSM	Modification	Relative abundance (PSM/unit length)
[R].TADGDEGGK.[H]	PsbP	1		1.58
[K].YYGQTVSNINEVLAK.[L]	PsbQ	16		0.99
[K].LFSSIDNLDHAAK.[I]	PsbQ	56		
[R].AKVSASEILNVK.[Q]	PsbQ	12		
[K].VSASEILNVK.[Q]	PsbQ	22		
[R].KAWPSLQNDLR.[L]	PsbQ	17		
[K].AWPSLQNDLR.[L]	PsbQ	12		
[R].DGTLPYTK.[D]	PsbQ	24		
[K].SPTEAEKYYGQTVSNINEVLAK.[L]	PsbQ	1		
[R].FYLQPLPPTEAAQR.[A]	PsbQ	10		
[K].SLQELTSK.[L]	PsbQ	52		
[K].DRFYLQPLPPTEAAQR.[A]	PsbQ	6		
[K].SPTEAEK.[Y]	PsbQ	1		
[K].IKVDKPLGIGGGMK.[L]	PsbR	15	1xOxidation [M13]	0.89
[K].IKVDKPLGIGGGMK.[L]	PsbR	23		
[K].VDKPLGIGGGMK.[L]	PsbR	23		
[K].VDKPLGIGGGMK.[L]	PsbR	42	1xOxidation [M11]	
[R].KPTGKGVYQFVDK.[Y]	PsbR	3		
[K].GVYQFVDK.[Y]	PsbR	18		
[R].FVDEPTTGLEK.[A]	PsbS	10		0.06
[K].GPLFGFTK.[S]	PsbS	3		
[K].VEDGLFGTSGGIGFTK.[E]	PsbS	1		
[K].SNELFVGR.[L]	PsbS	2		

Table 2.5 continued.

Peptide	Subunit	PSM	Modification	Relative abundance (PSM/unit length)
[K].TKGPLFGFTK.[S]	PsbS	1		0.06
[K].YAPVCVTMPSAR.[I]	PsbTn	6	1xOxidation [M8]; 1xCarbamidomethyl [C5]	0.14
[K].YAPVCVTMPSAR.[I]	PsbTn	5	1xCarbamidomethyl [C5]	
[K].KYAPVCVTMPSAR.[I]	PsbTn	2	1xCarbamidomethyl [C6]	
[K].KYAPVCVTMPSAR.[I]	PsbTn	1	1xOxidation [M9]; 1xCarbamidomethyl [C6]	
[K].YAPVCVTMPSAR.[I]	PsbX	6	1xOxidation [M8]; 1xCarbamidomethyl [C5]	0.12
[K].YAPVCVTMPSAR.[I]	PsbX	5	1xCarbamidomethyl [C5]	
[K].KYAPVCVTMPSAR.[I]	PsbX	2	1xCarbamidomethyl [C6]	
[K].KYAPVCVTMPSAR.[I]	PsbX	1	1xOxidation [M9]; 1xCarbamidomethyl [C6]	
QIAESPQEDNR	PsbY	4		0.02
[K].SDPDVADAVTELR.[A]	Psb27	6		0.33
[R].ILEEMDSVEK.[A]	Psb27	16	1xOxidation [M5]	
[K].NKSDPDVADAVTELR.[A]	Psb27	7		
[R].ILEEMDSVEK.[A]	Psb27	13		
[R].ATSNFWVAK.[Y]	Psb27	15		
[R].FDQPSVFDSSK.[E]	Psb28	5		0.03
[R].AQTSASLAEFSSR.[E]	Psb29	3		0.02
[R].EGEIEDVLK.[D]	Psb29	3		
[R].TAENFSNSTGEQGYPGGK.[F]	CP24	24		0.23
[K].KSWIPAVK.[G]	CP24	20		
[K].SWIPAVK.[G]	CP24	16		

Table 2.5 continued.

Peptide	Subunit	PSM	Modification	Relative abundance (PSM/unit length)
[R].WAMLGAAGCVIPEAFNK.[F]	CP26	8	1xCarbamidomethyl [C9]	0.62
[R].WAMLGAAGCVIPEAFNK.[F]	CP26	1	1xOxidation [M3]; 1xCarbamidomethyl [C9]	
[K].TGALLLDGNTLNYFGK.[N]	CP26	6		
[K].VAVASPADDELAK.[W]	CP26	42		
[K].FGANCGPEAVWFK.[T]	CP26	3	1xCarbamidomethyl [C5]	
[R].ITNGLDFEDK.[L]	CP26	19		
[K].DPDQAALLK.[V]	CP26	23		
[K].AKVAVASPADDELAK.[W]	CP26	13		
[K].YQAFELIHAR.[W]	CP26	34		
[K].LHPGGPFDPLGLAK.[D]	CP26	24		
[R].IFLPDGLLDR.[D]	CP26	4		
[K].KPEDFAK.[Y]	CP26	5		
[K].STSLQPYSEVFGLQR.[F]	CP29	6		0.26
[K].NLAGDIIGTR.[T]	CP29	24		
[K].SGFSTDRPLWYPGAK.[A]	CP29	11		
[K].KPILQLAEIK.[H]	CP29	8		
[R].FRECELIHGR.[W]	CP29	13	1xCarbamidomethyl [C4]	
[R].WAMLGALGCVFPELLAR.[N]	LHCII	1	1xOxidation [M3]; 1xCarbamidomethyl [C9]	0.99
[K].NVSSGSPWYGPDRVK.[Y]	LHCII	9		
[R].WAMLGALGCVFPELLAR.[N]	LHCII	1	1xCarbamidomethyl [C9]	
[K].NVSSGSPWYGPDR.[V]	LHCII	59		
[K].LGPTASEILSEGR.[V]	LHCII	1		

Table 2.5 continued.

Peptide	Subunit	PSM	Modification	Relative abundance (PSM/unit length)
[K].FGEAVWFK.[A]	LHCII	8		
[K].NRELEVIHCR.[W]	LHCII	28	1xCarbamidomethyl [C9]	
[R].ELEVIHCR.[W]	LHCII	41	1xCarbamidomethyl [C7]	
[K].TAGKPKNVSSGSPWYGPDR.[V]	LHCII	1		
[R].WAMLGALGCVFPELLAR.[N]	LHCII	1	1xOxidation [M3]; 1xCarbamidomethyl [C9]	
[R].WAMLGALGCVFPELLAR.[N]	LHCII	1	1xCarbamidomethyl [C9]	
[K].TVQSSSPWYGPDR.[V]	LHCII	25		0.99
[K].TVQSSSPWYGPDRVK.[Y]	LHCII	7		
[K].FGEAVWFK.[A]	LHCII	8		
[K].NRELEVIHCR.[W]	LHCII	28	1xCarbamidomethyl [C9]	
[R].ELEVIHCR.[W]	LHCII	41	1xCarbamidomethyl [C7]	
[K].LGPTASEIIGEGR.[I]	LHCII	1		
[K].TAGKPKTVQSSSPWYGPDR.[V]	LHCII	2		

2.3 Discussion

It is not clear why the OTG-PSII cores show low contrast in negative stain TEM when solubilized just in OTG detergent. Particles appear to have predominantly side-view orientations with or without glow-discharging the grids. Previous studies of membrane proteins use detergents, such as DDM, Brij-35, CHAPS, and digitonin both for negative stain and cryo-EM [76]. By seeing good contrast in the sucrose gradient C_2S_2 particles, which were solubilized in DDM, a detergent exchange for the OTG-PSII particles was the obvious next step. This exchange produced highly reproducible images with enough contrast for negative stain. These particles could be analyzed by SPA TEM and clearly show that the C_2 core complex exist in the purification. It was therefore important to test the activity and verify the subunit composition after these alterations were made. According to the data, modifications to create the OPDM sample did not have an effect on the activity nor on the complex composition.

In these studies, the purification method was chosen to improve on the oxygen evolution rates reported previously (as shown in Table 2.1). Controls for maintaining activity, such as minimizing light exposure and regulating temperature are common when working with photosystem II. However, the pH of the buffer was shown to be an important factor as well. When the pH is increased from 6.0 to 7.5, this causes a loss of extrinsic subunits from the complex as described by the reduction in oxygen evolving activity (Table 2.2). Excess calcium chloride was previously shown to restore oxygen evolving rates in PSII when extrinsic polypeptides have been depleted. This occurs because the calcium stabilizes the OEC and promotes electron transfer when the extrinsic

subunits have been removed [157]. Calcium chloride was added to show the restoration of oxygen rates was due to the loss of extrinsic subunits at pH 7.5 and that no other major changes occur.

Protein polishing steps, such as sucrose gradient ultracentrifugation or size exclusion chromatography, were eliminated from this purification because it was shown to reduce oxygen evolving activity rates. We reproduced the purification performed in recent structural methods of plant PSII (Figure 2.1 and 2.2) and found that the oxygen rates were severely lowered (Table 2.1). Part of the reason for the higher oxygen evolution in our OTG-PSII cores is due to the fact that there are more C_2 cores that are included in the assay. Because the amount of chlorophyll per assay does not distinguish between the chlorophyll in the antenna proteins versus chlorophyll from cores, there will be much more active, reaction centers per milligram of chlorophyll added into the assay than are added in the C_2S_2 or $C_2S_2M_2$ complexes. However, the chlorophyll *a/b* ratios of the C_2S_2 and $C_2S_2M_2$ complexes are 3:1 and 2.25:1, respectively. By accounting for the micromolar amount of chlorophylls per reaction center, we were able to calculate approximate, normalized, oxygen evolution rates for various types of PSII preparations (Table 2.3). The rates reported are lower than our active, C_2 cores. The stacked $C_2S_2M_2$ complex from the *Pisum sativum* structure comes close to having high activity rates compared to the OPDM sample. The BBY rates also have very high normalized rates, though the TEM imaging conditions for this type of sample have not been experimented with. It can be concluded that the choice and strategy for the protocol described here results in the highest activity of the PSII preparations to date.

At times the bands of the SDS-PAGE gels in the OTG-PSII and OPDM lanes appeared darker or more concentrated (Figure 2.3 and 2.4). This affect can be explained by the fact that chlorophyll is used to measure relative concentration of the protein. Because there is more overall chlorophyll in the BBY sample (due to the antenna proteins), less protein is required to achieve the 1 μg of chlorophyll for the assay. Thus, even though equal amounts of chlorophyll are being loaded into each lane, the resulting amount of protein is not exactly the same. The SDS-PAGE analysis still served the important purpose of identifying the major intrinsic and extrinsic subunits which are vital for oxygen evolution. There were clear signs of antenna proteins in the sample, which can be seen by the additional bands in the gels and the detection of peptide fragments by mass spectrometry. This could raise concern because these results seem to indicate the PSII complex contains the antennae proteins, resulting in variations of the C_2S_2 or $C_2S_2M_2$ types of PSII. However, TEM micrographs showed the predominant particle type was the C_2 complex. Therefore, we were not concerned with the presence of residual LHCII, CP24, CP26 and CP29 because they were not bound to the predominant type of PSII particles.

The UV-vis absorption spectra also support the production of predominantly C_2 PSII complexes. The presence of chlorophyll *b* peaks and shoulders in the spectra of the BBY sample indicated the abundance of antenna proteins such as LHCII, CP29, CP26, and CP24. The OTG-PSII and the OPDM samples show clear reduction of chlorophyll *b*, suggesting a loss of these antenna proteins. Similarly, the types of particles identified in the sucrose gradient B7-9 bands corresponded to the C_2 cores of our purification, which show a reduction of chlorophyll *b* content.

The tandem mass spectrometry analysis of peptides of the OPDM sample supplemented our understanding of the subunits that were present in our highly active PSII sample. Due to the high sensitivity of mass spectrometry and the heterogenetiy of our sample, we expected to see peptides of subunits such as PsbS, PsbR, LHCII, CP24 CP26, and CP29. These subunits are not included in the C_2 core particles, but it is important to know that these subunits are present in our preparation because these components could give clues to the reason for our highly active preparation. The unit used for measuring the amount of spectra per peptides was the peptide spectral match (PSM). To get an idea of the relative abundance of the subunit based on the PSM's, the unit is divided by the length of the subunit. Most PSM/unit length units fell between 0.4-1.6. The smaller intrinsic subunits such as PsbX were likely less abundant or difficult to detect due to its small size. PsbO had a PSM/unit length of 3.64 which could indicate that our sample has very high levels of PsbO or that the peptides are easily detected by mass spectrometry. Overall, the mass spectrometry results indicate the presence of the critical subunits of the C_2 PSII complex which is important for model building. The other subunits not associated to the core complex are likely in the OPDM sample but are either part of larger particles such as the C_2S_2 particles, or were co purified in the final sample preparation.

2.4 Experimental Procedures

2.4.1 Preparation of highly active, OTG-PSII cores

PSII samples were purified from market spinach. Leaves were washed and ground in a blender while in a cold environment, and the broken cells were suspended in a grind buffer (50 mM HEPES-NaOH, 400mM NaCl, 2mM $MgCl_2 \bullet 6H_2O$, 1mM EGTA, 2 mg/mL BSA, pH 7.5). Slurry was filtered of larger plant matter by squeezing through cheesecloth. After one centrifugation, the sample was washed and pelleted again to remove additional contaminants. The washed pellet was then solubilized with Triton X-100 (Sigma. T9284-500ML. Lot# 053K00262V), yielding PSII-rich thylakoid membrane fragments, called the BBY sample [57]. After a few washes in SMN buffer (400 mM sucrose, 50 mM MES-NaOH, 15 mM NaCl, pH 6.0), the membranes were then solubilized using, 0.8% octyl β-D-thioglucopyranoside (OTG) to achieve a final concentration of 0.4% of this detergent. After centrifugation to remove unsolubilized material, magnesium chloride solution is mixed with the sample at 10 mM. After a short incubation and centrifugation, much of the LHCII complexes are pelleted. Then, adding a PEG-6000 buffer to the supernatant at a 1:1 mixture, the sample is easily pelleted and resuspended in SMN buffer to an appropriate concentration, called the OTG-PSII cores [152]. Typically the sample is brought to 1 mg/mL of chlorophyll, flash frozen, and stored in a -70°C freezer.

2.4.2 Preparation of C_2S_2 supercomplexes

BBY method was performed as described previously [57]. C_2S_2 supercomplexes were made by adapting a previous method [50]. Briefly, BBY samples were thawed and

pelleted by spinning at 50,000*g for 10 minutes. Then, the sample was resuspended in SHN buffer (400 mM sucrose, 10 mM HEPES, 10 mM NaCl, pH 7.5) to 1 mg/mL. Next, an equal volume of detergent resuspension buffer (0.6% β-DDM, 10 mM HEPES, 10 mM NaCl, pH 7.5) was added and incubated for 1 minute, vortexing periodically. Sucrose gradients were made directly in the tube by freezing sucrose gradient buffer (650 mM sucrose, 0.008% β-DDM, 10 mM HEPES, pH 7.5) at -80°C and thawing at 4°C. On top of the sucrose gradients, 500 µg of chlorophyll were added gently. The tubes were placed into a SW-41 rotor, and spun at 247,600*g for 17 hours at 4°C to fractionate the sample into various components. Upon removal, bands were separated and analyzed by TEM and UV-vis spectroscopy.

2.4.3 Preparation of OPDM for TEM imaging

Aliquots of PSII at 1 mg/mL are thawed on ice. Centrifugation at 50,000*g for 15 minutes at 4°C allows the OTG-PSII cores to be pelleted. After removing supernatant, the pellet is resuspended in OPDM buffer (0.3% β-DDM, 50 mM MES-NaOH, 10 mM NaCl, and 20% sucrose, pH 6.0). The pellet is homogenized thoroughly and brought to a concentration appropriate for negative stain screening or cryo-EM imaging. The sample at this point is called OPDM.

2.4.4 Urea SDS-PAGE

Urea SDS-PAGE analysis was performed as previously described [158]. For PSII samples, 1 µg of chlorophyll was loaded per lane. Sample was run at 100 volts for 1.5 hours and then 150 volts for 45 minutes or until the chlorophyll and bromophenol dye ran

off the gel. Gels were stained with 0.05% Coomassie Brilliant Blue R in order to detect the subunits of PSII.

2.4.5 UV-vis absorption spectroscopy

10 μL of each type of sample was diluted with 990 μL of appropriate buffer (sucrose gradient buffer for sucrose gradient bands; OPDM buffer for OPDM sample; SMN buffer for BBY and OTG-PSII samples). Samples were placed into 1 mL quartz cuvettes and absorption spectra was taken from 350-750 nm, medium speed, on a Shimadzu UV-1700 spectrophotometer in room temperature.

2.4.6 Oxygen evolution assay

Oxygen assays were performed using a Clark-type oxygen electrode. Freshly prepared solution of recrystallized 0.5 mM 2,6-dichlorobenzoquinone and 1 mM potassium ferricyanide ($K_3[Fe(CN)_6]$) were used as external electron acceptors. Assay buffer contained 50 mM MES, 20% sucrose, 15mM NaCl, and 0.3% DDM. When required, calcium and chloride concentrations were adjusted from 3 M buffered stock solution.

2.4.7 Mass spectrometry

Proteins were solubilized in 8M Urea in Phosphate buffered Saline. The proteins in each sample were reduced, alkylated and digested with trypsin according to the FASP protocol [159]. The peptides were analyzed by nano-LC-MS/MS (Thermo Fisher Scientific, Q Exactive Plus, GT Systems Mass Spectrometry Core Facility), and peptide identification was performed as previously described [160] with the following

modifications. Reverse phase chromatography was performed using an in-house packed column (30 cm long X 75 μm ID X 360 OD, Dr. Maisch GmbH ReproSil-Pur 120 C18-AQ 1.9 μm beads) and a 120 minute gradient. The Raw files were searched using the Mascot algorithm (ver. 2.5.1) against the UniProt Swiss Prot database (downloaded July 7, 2017) via Proteome Discoverer 2.1. Only peptide spectral matches with expectation value of less than 0.01 ("High Confidence") were used.

CHAPTER 3. PREPARATION OF THE PSII COMPLEX FOR NEGATIVE STAIN TEM AND CRYO-EM

3.1 Introduction

Once a protein has been purified, a key step for single particle cryo-EM involves screening the protein sample by negative stain TEM. Identifying optimal conditions is critical for obtaining a good concentration and distribution for image processing, and it is necessary for characterizing the sample and its structure at low resolution. Here, the characterization and optimization of both biochemistry and TEM grid preparation of negatively stained samples of highly active PSII is described.

Evidence in the literature of PSII contradicts the positioning of extrinsic subunits of plant PSII from recent structures [42, 46]. In order to understand the localization of these extrinsic subunits, experiments can be performed to partially or fully remove them. Biochemical modifications of our PSII cores (by treatment of either high urea or high salt concentrations) were made and imaged under the standard protocol defined in this work. The results will allow for future experiments to answer questions about the locations of the extrinsic subunits in a highly active complex.

Because of its larger size (250-300 kDa) as a dimeric core complex, PSII is an ideal candidate for single particle cryo-EM. Membrane proteins have a hydrophobic region that tends to promote a preferred orientation. When PSII is in the native membrane, the extrinsic subunits face the lumen and this is typically considered a side-view orientation. When viewing the protein from the stroma or lumen, the complex takes on a rectangular

shape, which is considered the top or bottom view. By altering the surface charge of the TEM grid, it was found that the PSII particle orientation could be controlled. Additional important factors that played a role in the best conditions for negative stain were the choice of detergent and the concentration of the protein.

Diluting the OTG-PSII sample in buffer with additional OTG (up to 0.2%) resulted in a good distribution of individual particles compared to the protein aggregation which was seen when imaging the sample by diluting with detergent-free buffer. However, OTG substantially reduced the contrast of the PSII particles. Furthermore, previous literature involving negative stain TEM of PSII had shown particles with much better contrast [15].

The protein concentration was another important factor that was optimized. As explained in Chapter 1, PSII particles need to be highly concentrated and at the same time, well distributed for image processing. By testing glow discharge conditions of TEM grids, it was possible to alter the surface of the carbon film to increase sample adherence at low protein concentration. By re-solubilizing the PSII in DDM, the contrast was dramatically increased. Thus, the combination of parameters described here, allowed for the preliminary characterization by negative stain which served as a foundation for cryo-EM.

3.2 Results

3.2.1 Negative stain screening of PSII single particles

First trials of PSII single particle imaging by negative stain involved testing the purified OTG-PSII without any modifications. The starting chlorophyll concentration for all the OTG-PSII sample was typically brought to 1 mg/mL and the sample was then diluted with detergent-free buffer. Through the empirical discovery of using 0.2% OTG buffer rather than detergent-free buffer, we were able to avoid protein aggregation on the TEM grid (Figure 3.1).

Table 3.1 describes the various glow discharge conditions that were tested in combination with different chlorophyll concentrations that gave a good distribution of particles on the grid. The data indicated that the glow discharge conditions resulted in a mixture of both top and side-view orientations. Dilutions of 1:50 and 1:80 showed good particle distribution at low concentrations on the grid. A concentration of 1:20 resulted in a large number of particles for data collection. The dilutions were made by mixing the 1 mg/mL sample in SMN buffer which contained 0.2% OTG. The oxygen rates were tested after re-solubilization and did not result in a drop in activity (Chapter 2, Table 2.1).

Even with the glow discharge method, the OTG-PSII particles tended to have a preferred orientation. Particles on a non-glow discharged grid typically adhered to the carbon surface on their hydrophobic face, resulting in a side-view orientation. Particles on a glow discharged grid, adhered to the carbon on their hydrophilic face, resulting in a top or bottom-view orientation. Occasionally, extrinsic subunits could be identified in the side-view, but it was not clearly defined due to the poor contrast. Figure 3.1 shows a

representative TEM images of the OTG-PSII particles and their associated class averages (produced in EMAN 2.2). For negative stain image processing, it was faster and more efficient to use EMAN software. Upon 2D averaging, the particles in their side-view typically did not show signs of extrinsic subunits, which extend about 3 nm from membrane. However, some small densities could be interpreted from the averages.

Table 3.1 Glow discharge results from various concentrations and durations.

Dilution of 1 mg/mL of chl of OTG-PSII	Glow discharge duration	Orientation of particles	Distribution of particles
1:10	8 sec	top and side views	good
1:20	8 sec	top and side views	good
1:50	30 sec	side view	poor
1:50	15 sec	side view	fair
1:80	30 sec	top and side views	poor

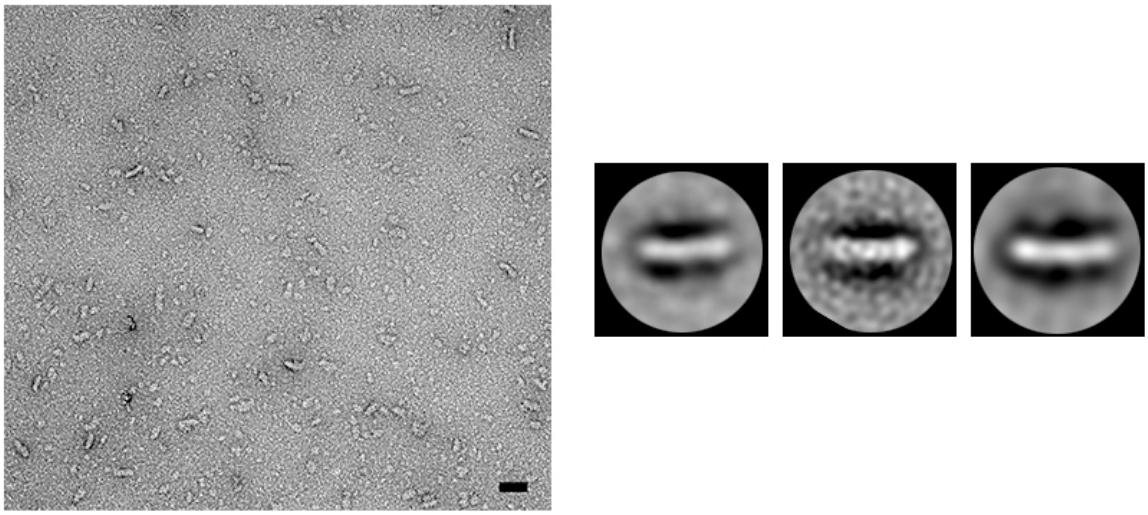


Figure 3.1 OTG-PSII particles in negative stain (2% uranyl acetate). Representative image of the particles seen on the TEM grid. Scale bar represents 50 nm. Three representative 2D class averages are shown to the right, illustrating the OTG-PSII predominant side-view orientation. Box sizes of 2D classes correspond to 750 Å.

The image processing step in negative stain was refined in order to improve the particles that were manually picked. By eliminating unsuitable particles (e.g. low contrast, damaged, improper size, etc.) that were contributing to noise in the data, the SNR could be boosted. Approximately 1000 particles were manually selected and reduced to approximately 500 good particles. From this set of 500 particles, several classes were made, which categorized PSII based on size (Figure 3.2). A total of 126 particles were measured to characterize PSII complex sizes present in this sample. Of these particles, about 30% were separated into a "small" class set (25 nm), 40% were separated into "medium" class set (32 nm), and the remaining particles were sorted into

"large" class set class (35 nm). Each of these sets were run through separate, reference-free 2D class averaging rounds to improve signal and avoid the averaging of particles of the three different sizes. three types of sizes. Figure 3.2 shows the resulting data. All averages appeared to show an obvious side-view orientation of the PSII particle, but the extrinsic subunits were not clearly seen.

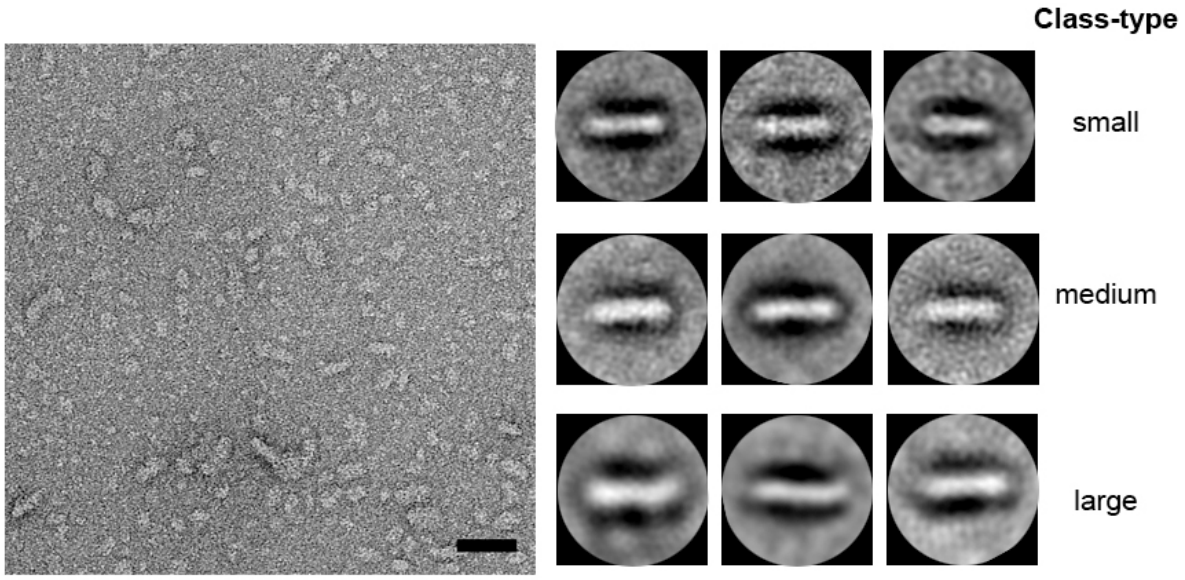


Figure 3.2 OTG-PSII class averages. The left image shows a representative micrograph used for particle picking. Scale bar corresponds to 50 nm. Three sets of averages (small, medium, and large) were generated in EMAN. All showed predominantly side-view orientation. Box sizes of 2D classes correspond to 750 Å.

After understanding how the OTG-PSII particles behave in negative stain TEM, the next task was to improve the image contrast. The decision for re-solubilizing the OTG-PSII in DDM buffer was made after replicating a sucrose gradient purification of PSII. Because this purification uses DDM to solubilize the protein and had good contrast,

we decided to test if this detergent could be used with our OTG-PSII cores in order to increase the contrast for TEM. The improvement in contrast and visibility of the particles was coupled with an adjustment to uranyl acetate staining protocol modified from previous work on single particle TEM preparations [76]. It was clear in the PSII sucrose gradient purification (band 9 and band 10) that the particles had a mixture of orientations and good particle distribution and contrast (Figre 3.3 A and B). The averages show mostly C_2S_2 type supercomplexes in two orientations with some density for the extrinsic subunits. The sample from band 9 of the sucrose gradient purification (Figure 3.3 A) resulted in 2D averages of C_2S_2, C_2S, and C_2 particles that may have occurred due to subunits being lost during the purification.

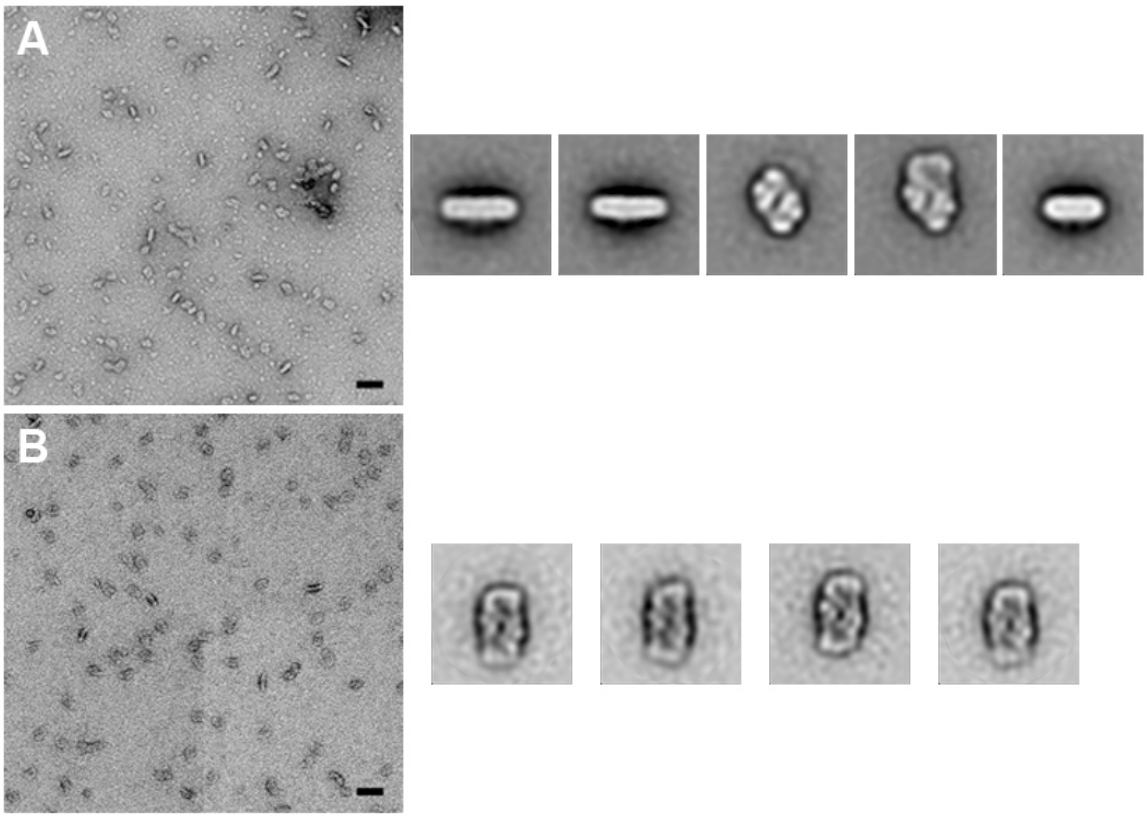


Figure 3.3 Sucrose gradient purification of PSII complexes with associated 2D class averages. (A) shows particles extracted from band 9, and (B) shows particles extracted from band 10 of the sucrose gradient. Scale bars represent 50 nm. To the right of each of the micrographs are the representative class averages. Box sizes for 2D classes correspond to 580 Å.

After examining the results of the sucrose gradient purification, the OTG-PSII cores were re-solubilized in β-DDM to create the OPDM sample. The particles were diluted to an optimal 1:20 concentration and applied to a glow discharged grid. The images showed particles oriented predominantly in their top or bottom views, with few side-view orientations. Even before image processing, it was apparent that these OPDM core particles were likely the C_2 complexes because of their overall size, shape and characteristic dark-center, which forms from uranyl acetate pooling between the two monomers of PSII. Figure 3.4 shows a comparison of PSII particles from band 5 of the sucrose gradient purification (which contained C_2 core dimers of PSII) and OPDM PSII particles at pH 6.0. Both particles have high contrast and show predominantly top/bottom orientations. Data was collected to generate 2D class averages of the OPDM PSII, which confirmed that these particles were likely the C_2 core complex (Figure 3.5) due to their 10 x 20 nm dimensions and overall parallelogram shape.

The OPDM C_2 core 2D average was compared to the sucrose gradient C_2S_2 supercomplex average (Figure 3.6). A reference model from the PDB: 3JCU structure was used for the comparison. The core dimer, made of the intrinsic and extrinsic subunits, has a length of approximately 20 nm and a width of 10 nm. The C_2S_2 supercomplex has a length of about 30 nm and a width of 16 nm. There are well-defined densities for each monomer of PSII in the negative stain OPDM C_2 core 2D average which correspond to the monomer in the 3JCU structure (Figure 3.6 A and B). Additionally, the C_2 core complex of PSII can be seen within the C_2S_2 supercomplex (Figure 3.6 C and D). Although not as strongly defined, the densities of the heterotrimeric

extrinsic subunits (PsbO, P, and Q) can be outlined in the negative stain average. These densities align with the positioning of PsbO, P and Q in the 3JCU structure.

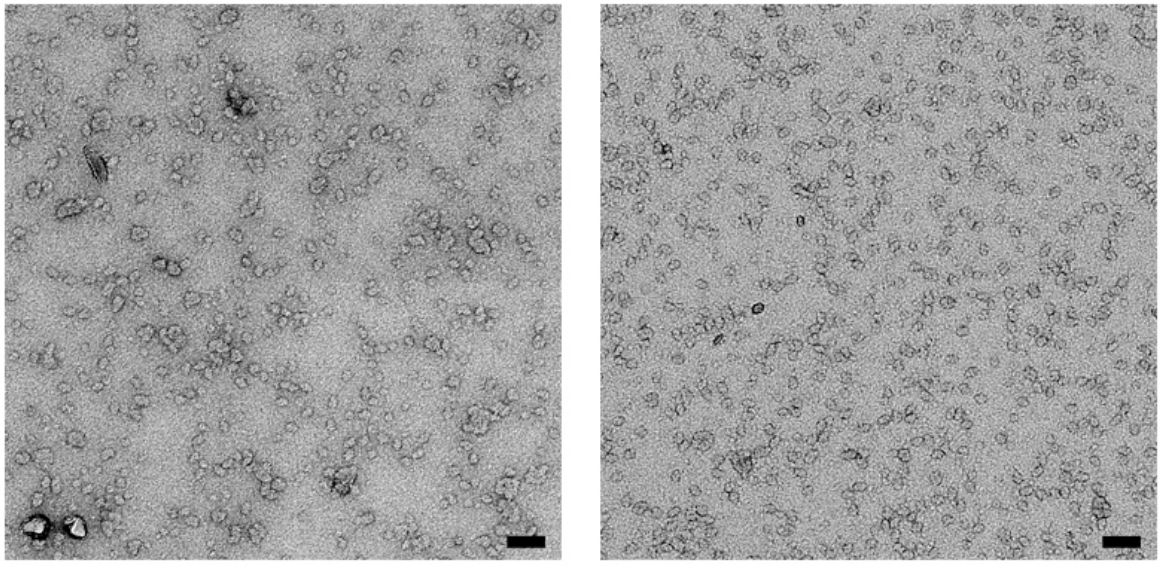


Figure 3.4 Comparison of sucrose gradient PSII particles and the OPDM PSII by negative stain. The micrograph on the left is from the extraction of band 5 from the sucrose gradient. The micrograph on the right is from the OPDM PSII sample. Scale bars correspond to 50 nm.

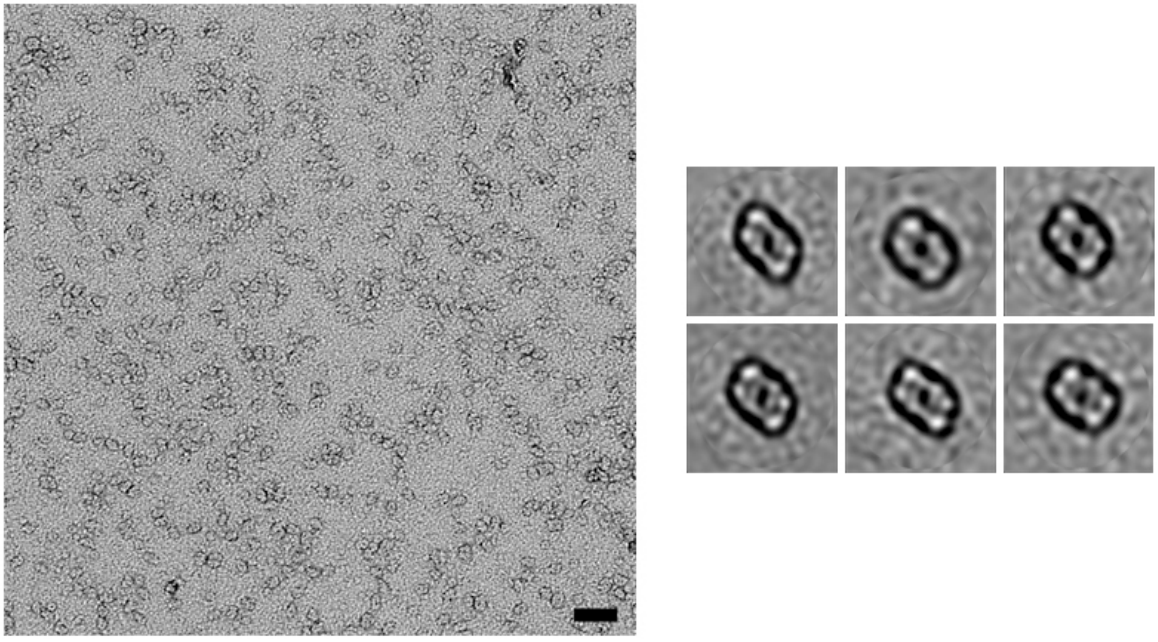


Figure 3.5 OPDM PSII with associated 2D class averages. Representative image of particles found in the OPDM preparation. Scale bar represents 50 nm. The 2D class averages generated from these images are depicted to the right. These averages strongly support that the particles are the C_2 type PSII. Box sizes of the 2D classes represent 580 Å.

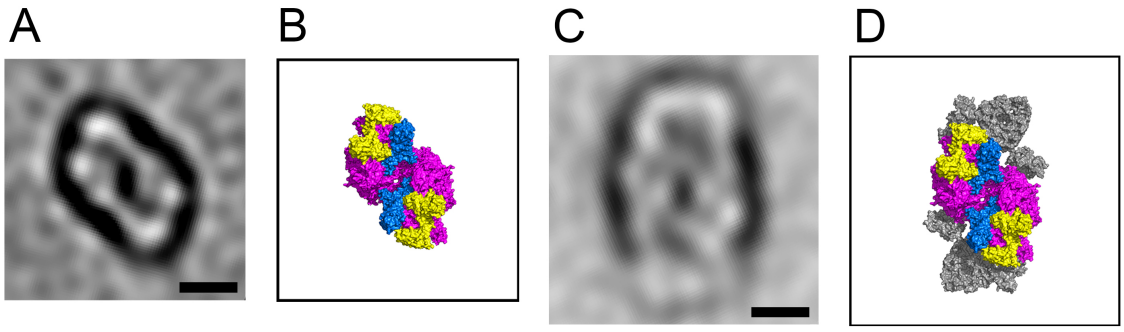


Figure 3.6 Negative stain 2D class averages from two purifications of PSII compared to the PDB: 3JCU structure [18]. (A) OPDM C_2 type PSII. (B) The 3JCU structure with only the intrinsic subunits (magenta) and three major extrinsic subunits (PsbO, blue; PsbP and PsbQ, yellow). (C) sucrose gradient C_2S_2 type PSII extracted from band 9. (D) The 3JCU with the core proteins plus antenna proteins: LHCII, CP26 and CP29 (grey). The particles of the OPDM sample strongly support the idea that we have the C_2-type PSII. Scale bars correspond to 10 nm.

3.2.1.1 Salt-wash and Urea-wash OPDM

Experiments to determine the localization of the extrinsic subunits on our active PSII preparation can be performed by removing the extrinsic subunits. Initial testing was performed to partially remove PsbP and PsbQ by high salt treatment. Additionally, a urea wash treatment was also performed for the removal of all three extrinsic subunits. Table 3.2 summarizes each of the samples that were prepared and the combination of extrinsic subunits present. The SDS-PAGE analysis shows that high salt wash treatment removes PsbP and PsbQ specifically, while the urea wash treatment removes all three extrinsic subunits (Figure 3.7).

The sample was negatively stained under the same conditions as previously described, using a 1:400 dilution when staining on a glow discharged grid, and using a 1:20 dilution when staining on a grid that was not glow discharged. The reason for the change in dilutions with or without glow discharge can be attributed to the fact that glow discharging a TEM grid causes the carbon surface to attract more particles, therefore less concentrated samples are needed. With glow discharge, the particles adopted a predominant top-view orientation, where the approximate dimensions were 10 nm by 20 nm for the control OPDM PSII (Figure 3.8 A) salt wash PSII (Figure 3.9 A), and urea wash PSII (Figure 3.10 A). The averages of these particles typically had a parallelogram shape. Without glow discharge, the particles were predominantly in the side-view orientation, providing strong contrast, but no clear extrinsic subunit definition. The control OPDM PSII is shown in its side view in Figure 3.8 B. The salt wash PSII (Figure 3.9 B) averages had some stacked C_2 particles that formed a dimer of dimers with extrinsic subunits facing outward. This stacking effect in PSII dimers occurs when there

are a high level of cations in the buffer and has been suggested to be a mechanism of how the thylakoid membranes facilitate stacking to form grana [60, 161]. The urea wash PSII averages (Figure 3.10 B) did not show signs of extrinsic subunits. The overall contrast of the top-view and side-view of urea wash PSII was not as good compared to the control OPDM sample. The particles also seemed to lose their characteristic, well-defined parallelogram shape.

Table 3.2 Summary of the PSII samples to understand the localization of the extrinsic subunits.

Sample preparation	Major extrinsic subunits
Control OPDM PSII	PsbO, PsbP, PsbQ
Salt wash PSII	PsbO
Urea wash PSII	none

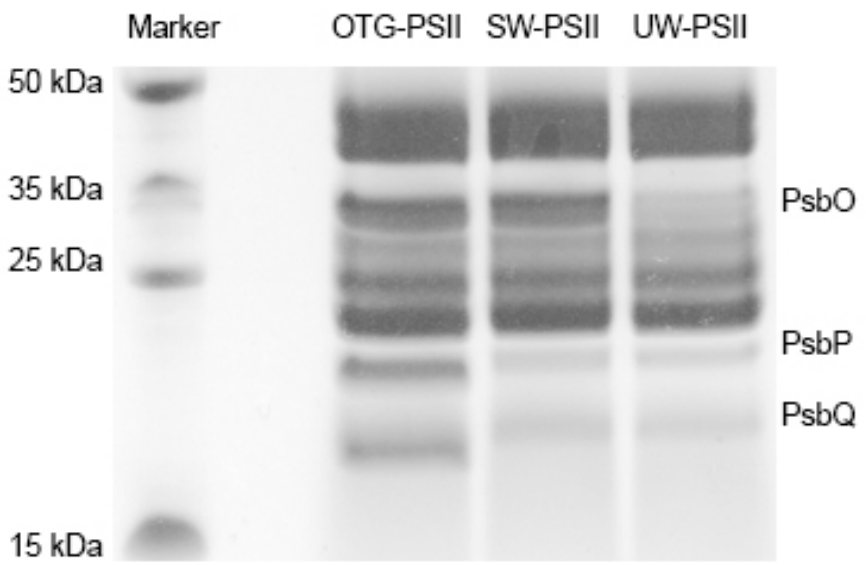
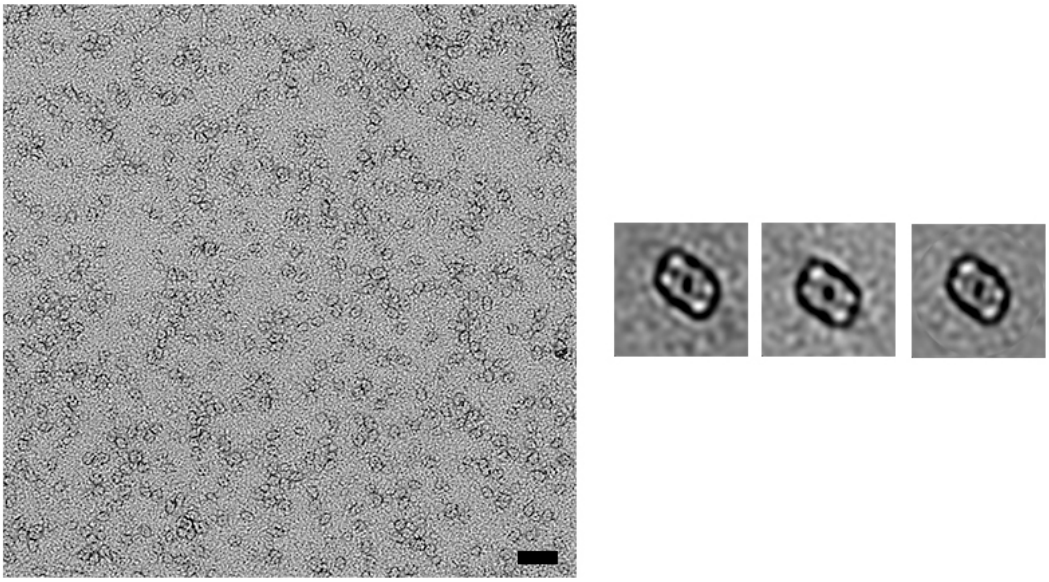


Figure 3.7 SDS-PAGE analysis of control OTG-PSII, PSII after salt wash treatment, and PSII after urea wash treatment. Fading of bands at each respective location indicate the removal of the majority of the extrinsic subunits from the complex.

A



B

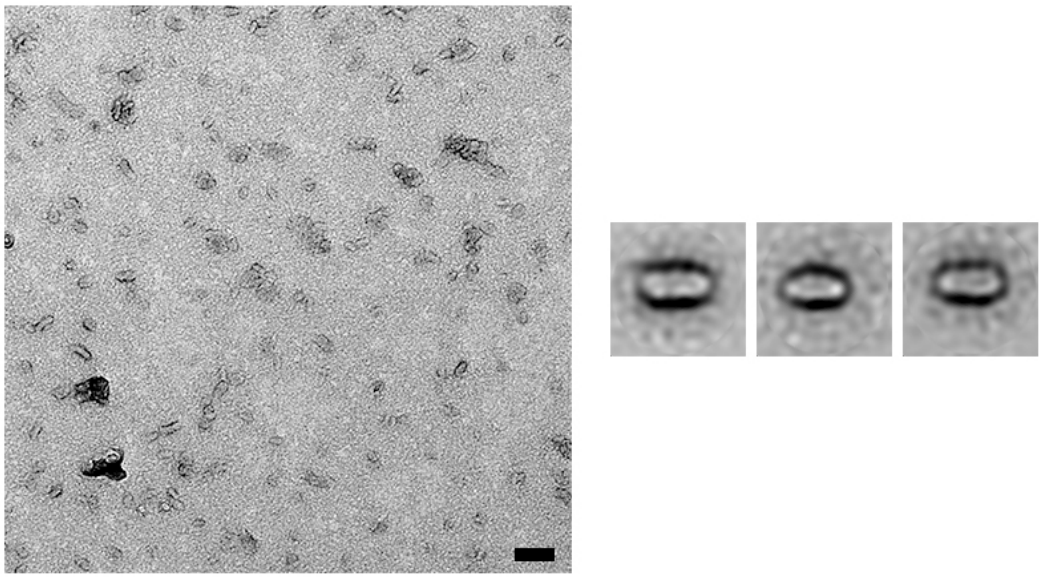


Figure 3.8 OPDM PSII control samples imaged by negative stain TEM. Representative micrographs are shown on the left of a glow discharged grid (A) and a non-glow discharged grid (B). Scale bars represent 50 nm. Representative 2D averages of each type are shown to the right. Box sizes correspond to 580 Å.

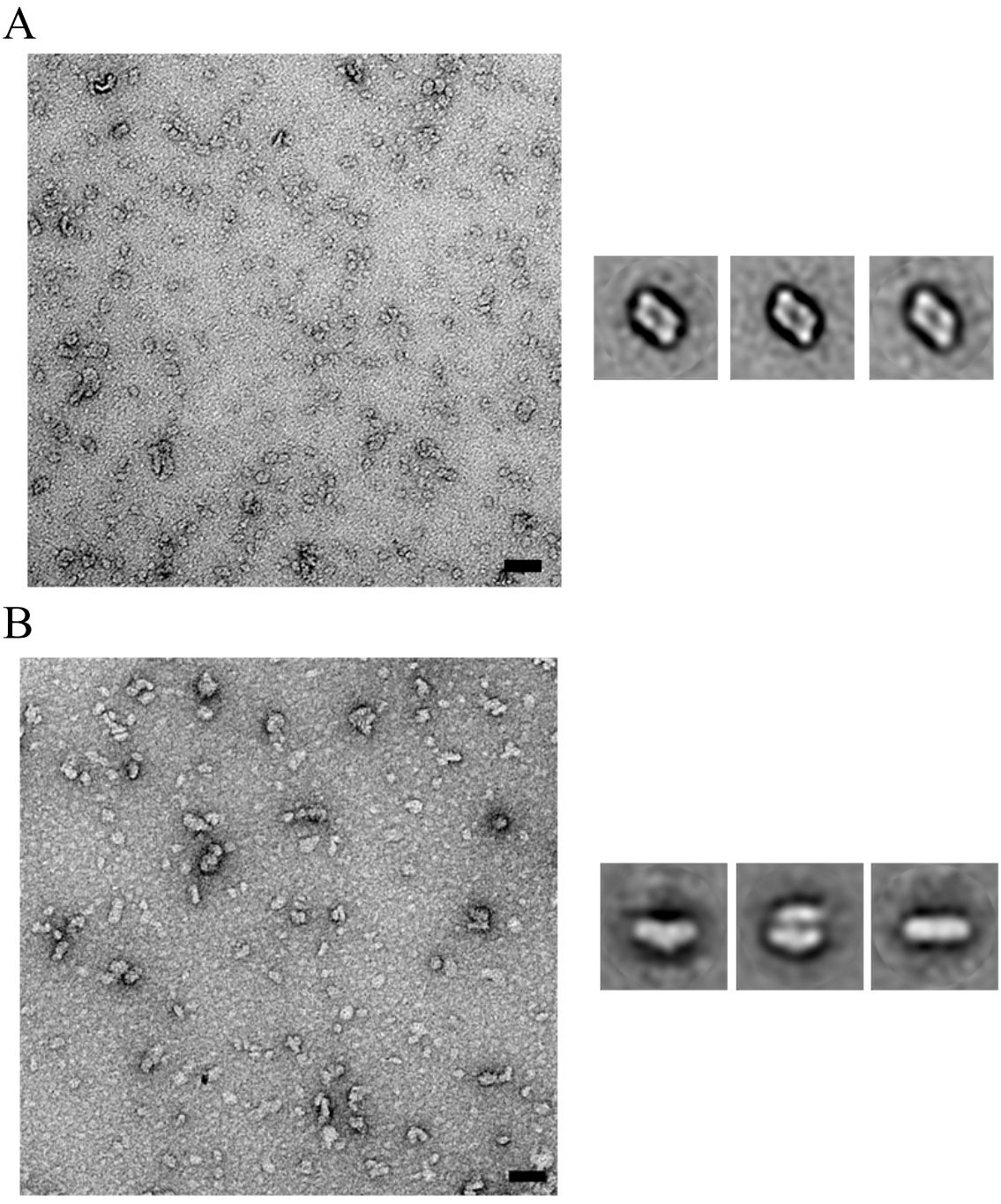


Figure 3.9 Salt wash treatment of OPDM PSII. Representative micrographs are shown on the left of a glow discharged grid (A) and a non-glow discharged grid (B). Scale bars represent 50 nm. Representative 2D averages of each type are shown to the right. Box sizes represent 580 Å.

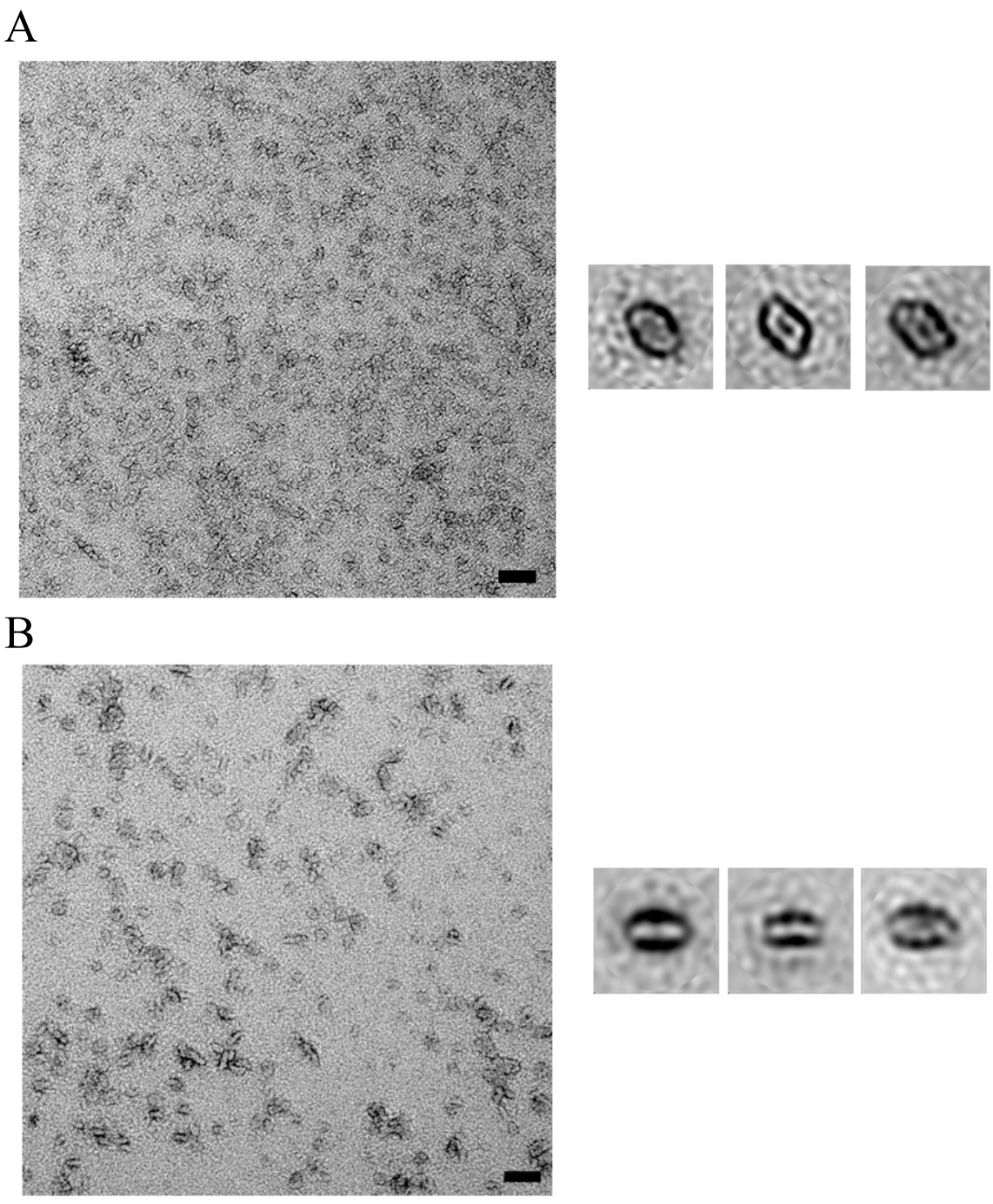


Figure 3.10 Urea wash treatment of OPDM PSII. Representative micrographs are shown on the left of a glow discharged grid (A) and a non-glow discharged grid (B). Scale bars represent 50 nm. Representative 2D averages of each type are shown to the right. Box sizes represent 580 Å.

3.2.2 Cryo-EM imaging and data collection

Conditions were tested for preparing the OPDM sample for cryo-EM. Various blot times and concentrations were tested and are shown in Table 3.3. Initially, the active OTG-PSII and the sucrose gradient samples were tested. Four second blot times showed good or fair ice quality. This qualitative description refers to the level of thickness of the ice. If the ice is too thick, the quality is deemed, poor. If there are certain areas of the grid that have decent thickness and can be imaged on, then the quality is considered fair. If most of the grid can be used for data collection due to thin, vitreous ice, the quality is considered good.

The sucrose gradient PSII was tested in cryo-EM to replicate conditions which were known to work based on previous studies [18]. Figure 3.11 shows a cryo-EM image of the sucrose gradient PSII sample at 0.5 mg/mL of chlorophyll. The results of the sucrose gradient cryo-EM data informed the OPDM PSII grid preparation. Three chlorophyll concentrations were tested: 1, 2, and 3 mg/mL. The concentrations at 2 and 3 mg/mL were determined to be too highly concentration for cryo-EM grid preparation. The OPDM sample at 1 mg/mL produced the best ice thickness and showed clear signs of particles around 20 nm in length in the holes of the perforated carbon film (Figure 3.12). Although the particle concentration in the holes was not as high as expected (average of ~5-6 particles per image), approximately 2000 images were collected.

Table 3.3 Cryo-EM conditions tested at Titan Krios for various PSII samples.

PSII type	Chlorophyll concentration (mg/mL)	Blot time (sec)	Ice quality
OTG-PSII	0.4	4	poor
OTG-PSII	0.4	4	fair
OTG-PSII	0.4	4	fair
Sucrose gradient PSII	0.5	4	good
Sucrose gradient PSII	0.5	4	good
Sucrose gradient PSII	0.5	5	good
OPDM	1	5	good
OPDM	2	5	poor
OPDM	3	4	poor
OPDM	3	5	poor

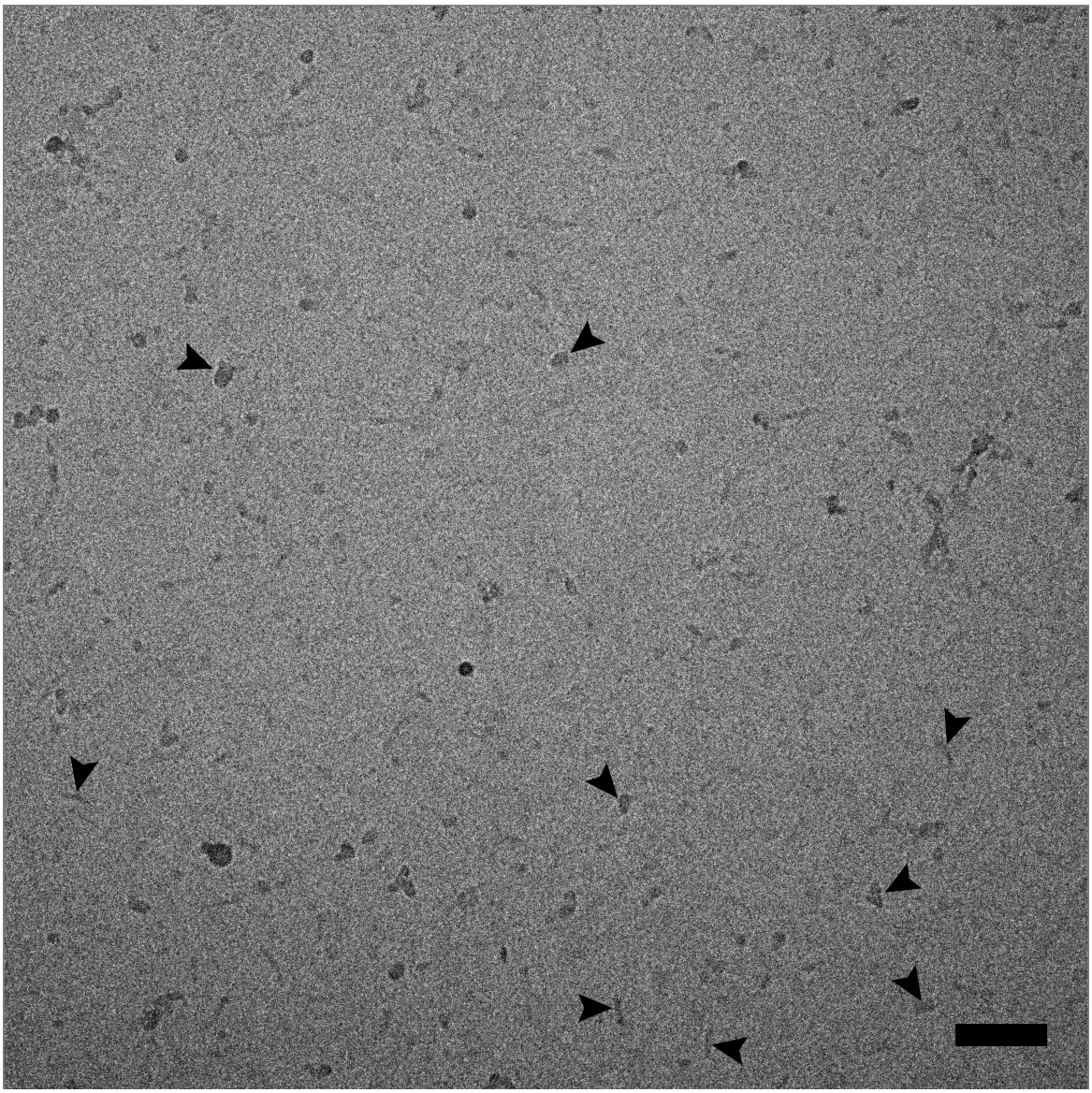


Figure 3.11 Cryo-EM image of inactive sucrose gradient PSII. Black arrowheads identify some of the ideal particles that would be picked in image processing. Scale bar represents 100 nm.

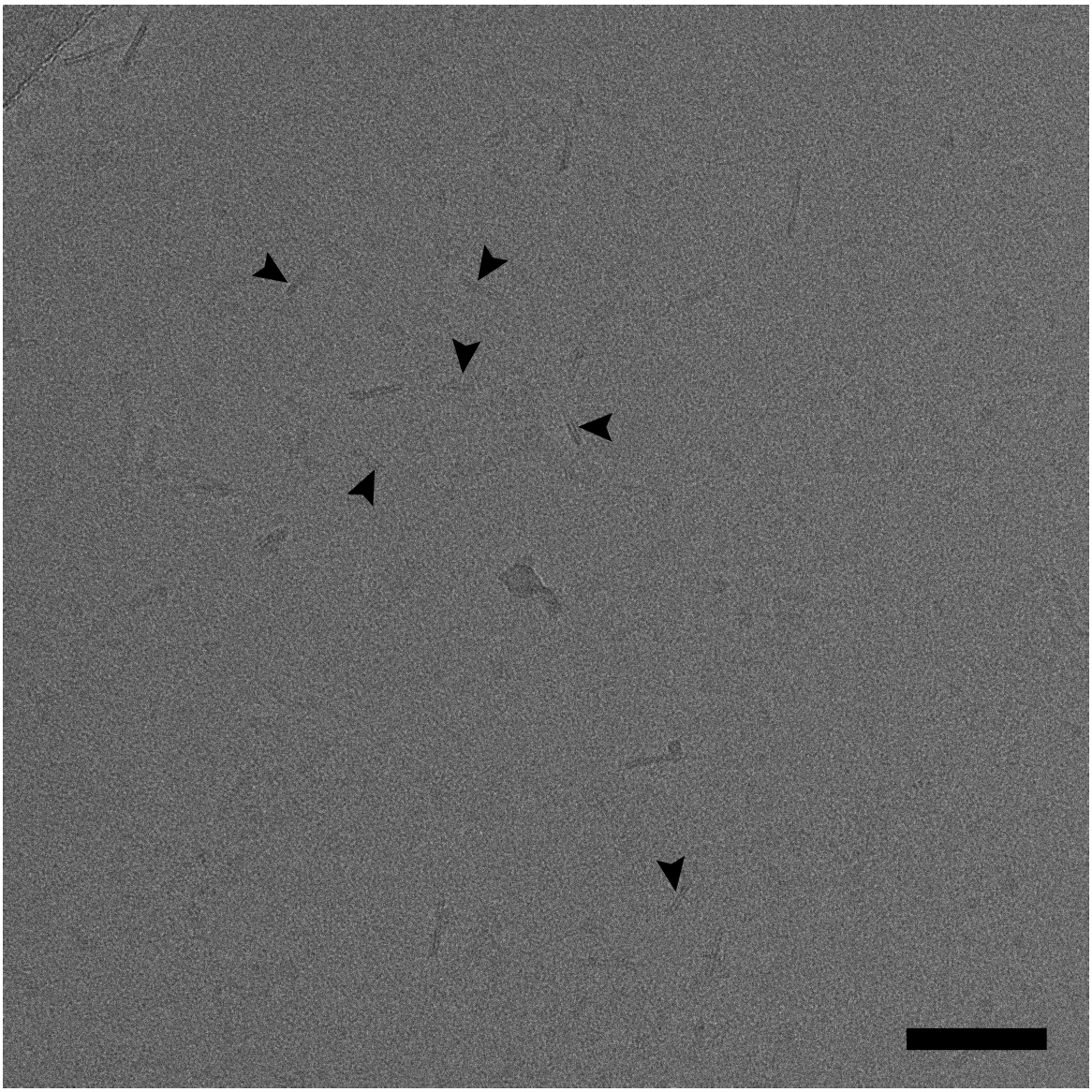


Figure 3.12 Cryo-EM image of the OPDM PSII sample. PSII particles are indicated by black arrowheads. Scale bar corresponds to 100 nm.

After manually sorting the best images and performing motion correction and CTF estimation, approximately 1800 images were selected and 53,742 particles were picked from these images. The particle coordinates were extracted and the data was binned by 2, to increase the contrast, decrease the total file size of the particle stack, and allow the image processing to run faster. The resulting particle images had an angstrom per pixel size of 2.02, resulting in a total box size of 320 Å. Several processing packages were used, such as Relion, cisTEM, FREALIGN, and cryoSPARC. Testing various software packages with our data was important for identifying the best approach for 3D reconstruction. CisTEM produced the best 3D reconstructed models. In Figure 3.13 a few of the 2D class averages show clear signs of the C_2 PSII. The extrinsic subunit density can be seen in side-view orientations which could not be seen as well in the negative stain data. Details of the 3D reconstruction results and interpretation are in the next chapter.

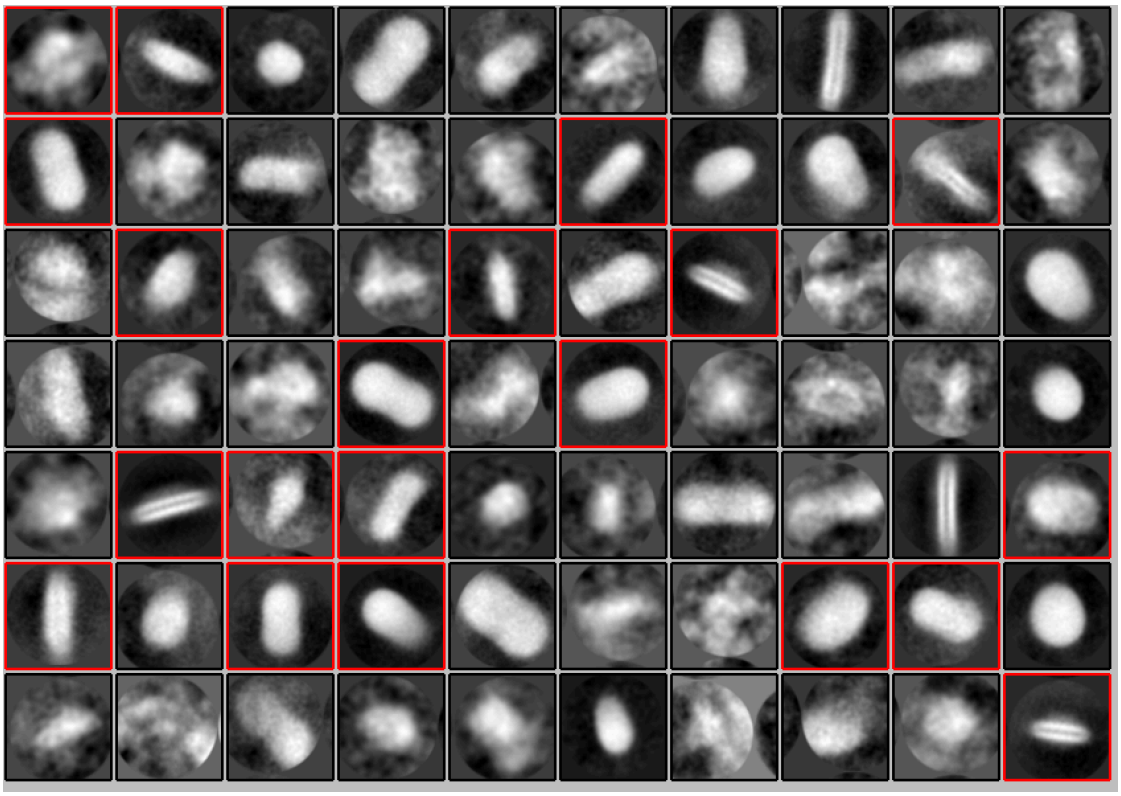


Figure 3.13 Representative 2D class averages of OPDM PSII. The red boxes indicate the averages that were used for 3D refinement. Box sizes correspond to 320 Å.

3.3 Discussion

The successful grid preparation was a critical step in the PSII project. Not only did these rigorous experiments pave the way for the best cryo-EM conditions for PSII, they also allowed for reproducible future experiments where detailed structure-function studies of the PSII complex can be performed under these reproducible conditions. The method of purifying a highly active PSII is not new. Typically, this purification protocol is used for biochemical assays and spectroscopic measurements. In structural biology, it is much more common to purify the protein so that it is monodisperse. This is typically done by size exclusion chromatography or sucrose gradient centrifugation. This has advantages in both cryo-EM and crystallography. In the case of PSII, previous structural models were obtained using the sucrose gradient protocol, resulting in a complete loss or reduction of activity.

Cryo-EM of the OTG-PSII cores required multiple modifications to identify a buffer that maintained activity and allowed good visualization by TEM. After testing several concentrations, an optimal chlorophyll concentration was reached at around 0.05 mg/mL, which was a 1:20 dilution of the original starting concentration. At this concentration we began to see particles with a good distribution, however the method was not very reproducible and would yield images that could not easily be processed. Particles from this preparation also had a predominant side-view orientation which did not show clear extrinsic subunits and the contrast was low. By manually selecting particles and sorting the data into three classes based on size, it was shown that the particles still did not show clear features of extrinsic subunits (Figure 3.1). It is not

known why exactly the OTG detergent was not optimal for TEM imaging. However, by imaging the sucrose gradient PSII, which was solubilized in DDM, it was clear that the difference in detergent improved the contrast and reproducibility in negative stain (Figure 3.3). The modification in the staining protocol also had an impact as the glow discharge produced top-view orientations of the particles.

With the new detergent and negative stain information, the OTG-PSII cores were exchanged into DDM to make the OPDM sample. The images were reproducible, and the particles were easy to distinguish. This in turn made the image processing significantly more efficient in the manual particle picking step. The 2D class averages yielded results that lead us to think that we had the C_2 type PSII. The glow discharge method gives the particles a predominant top-view orientation, which makes it difficult to determine extrinsic subunit presence, but helped identify our particles as the C_2 type PSII. We were confident that the extrinsic subunits were still intact due to the high oxygen evolution rates and the SDS-PAGE analysis. The presence of the extrinsic subunits was confirmed in cryo-EM 2D averaging, where there is clear density for these subunits (Figure 3.13).

The high reproducibility of the OPDM sample preparation allowed for the experiments which could aid in determining the extrinsic subunit positions. OPDM PSII cores were treated with a high salt treatment, which removes the PsbP and PsbQ subunits. The negative stain images showed that it was possible to get top-view orientations and these 2D averages looked similar to the averages from the OPDM control (Figure 3.7, 3.8, and 3.9). This suggests that the overall structure of the C_2 core does not change significantly, despite the partial or complete loss of extrinsic subunits. By using grids that were not glow discharged, we were able to get a better mixture of side and top view

orientations, though it was difficult to determine whether or not the extrinsic subunits were present. Negative stain averages of PSII usually do not produce high resolution detail, however that level of detail was apparent in cryo-EM, as seen in the 2D cryo-EM averages of the OPDM PSII (Figure 3.13). By removing the extrinsic subunits, it will be possible to obtain a structure of PSII and properly assign the locations of the extrinsic subunits.

3.4 Experimental Procedures

3.4.1 Removal of PSII extrinsic subunits

The OTG-PSII samples were depleted of extrinsic subunits following a previous protocol [41]. Samples were incubated in high ionic strength buffer (400 mM sucrose, 50 mM MES-NaOH,, and 2 M NaCl, pH 6.0) to extract the extrinsic polypeptides, PsbP and PsbQ. To remove PsbO, the sample was instead incubated with concentrated urea buffer (400 mM sucrose, 50 mM MES-NaOH, and 2.6 M urea, 200 mM NaCl, pH 6.0). The urea washed PSII and the salt washed PSII were buffer exchanged in OPDM buffer before imaging by TEM.

3.4.2 Grid preparation for OTG-PSII single particle trials

Grids were prepared as described in [162]. Briefly, 400 mesh copper TEM grids were coated with a thin layer of carbon film. The film was prepared by evaporating a graphite rod onto a freshly cleaved piece of mica using a Cressington carbon coater 108carbon/A. The carbon layer was then floated onto water and dropped carefully over the copper grids and left to dry. 2 μL of purified OTG-PSII sample is added onto the carbon coated grid and incubated for 1 minute. Next, the sample was blotted from the side with a piece of Whatman 4 filter paper. 2 μL of 1% or 2% uranyl acetate were added to the grid and incubated for 30 seconds before blotting away.

3.4.3 Grid preparation for OPDM single particles

Carbon coated grids were prepared as described in the previous section (3.4.2). The major difference was the grids were glow discharged at 15 mA for 15 seconds, giving the

carbon surface a hydrophilic nature. To obtain a mixture of particle orientations, the glow discharge step was removed. The grid was prepared as described in [76] with some modifications. Briefly, 2 drops of water and 1 drop of 2% uranyl acetate (50 µL each) were placed on a piece of parafilm. A freshly glow-discharged grid was held with a pair of anti-capillary tweezers and 4 µL of OPDM sample was incubated on the grid for 2 minutes. The grid was then placed onto the first droplet of water and held carefully at the surface for 10 seconds before being blotted with a piece of Whatman 1 filter paper. This was repeated for the next drop of water and uranyl acetate.

3.4.4 Screening for PSII by transmission electron microscopy

Grids were screened using a JEOL JEM-1400 transmission electron microscope operating at an accelerating voltage of 120 kV. This microscope was equipped with a side-mounted Orius SC1000 (Gatan) and a bottom-mounted 2k x 2k Ultrascan 1000 (Gatan). Single particles of PSII were first imaged at low magnification (5000x-8000x) with the Orius CCD camera. Areas of interest were saved during this step to be imaged later at higher magnification. For PSII single particles, the Ultrascan CCD camera was used to collect images at a nominal magnification of 30,000x for initial negative stain data collection. Defocus was determined by computing a fast Fourier power spectrum of an area of the image.

3.4.5 Image processing negative stain data for 2D classes

Image processing was performed using EMAN 2.2 software [134]. Images were collected at 30,000x magnification at a defocus range of 0.8 to 1.2 µm. Micrographs were imported into the program and particles were picked using interactive manual particle

picking (particle size of 85 pixels and box size of 125 pixels). CTF estimation and correction was performed in EMAN, using CTFFIND [163]. Particles were extracted and 2D class averages were generated.

3.4.6 Cryo-EM grid preparation

Quantifoil R 2/2 grids were used for cryo-EM. The grids were glow discharged in an Argon gas environment. 3 μL of PSII at 1 mg/mL of chlorophyll were applied on the grid and plunged into liquid ethane using a semi-automatic plunge device (Thermo Fisher Scientific, Vitrobot, Mark IV), with a blotting time of 4-6 seconds and a blot force of 2 at 100% humidity and at 4°C.

3.4.7 Cryo-EM carbon backed grid preparation

A Petri dish with water was used to float a thin layer of carbon. C-flat grids were submerged into the water and brought to the surface of the carbon layer, creating cryo-wells. The grids were baked at 50°C overnight. Grids were glow discharged and prepared for cryo-EM as described in the previous section (3.4.6).

3.4.8 Cryo-EM data collection

The images used for data collection were taken at Florida State University's Biological Science Imaging Resource on the Titan-Krios electron microscope (300 kV), equipped with a DE-64 direct electron detector. 32 movie frames were taken per image for motion correction. Approximately 2,000 images were collected at a calibrated magnification of 37,000 yielding a pixel size of 1.01 Å. The dose rate was approximately 60 $e^-/Å^2$ and the defocus was between 0.8 and 2.00 μm.

CHAPTER 4. CRYO-EM OF HIGHLY ACTIVE PSII

4.1 Introduction

After obtaining cryo-EM data of single particles, image processing is performed to generate a 3D reconstruction, and then the model can be interpreted. Defining a pipeline for processing data for a 3D reconstruction of PSII is critical since factors such as particle size, heterogeneity, symmetry, and image contrast quality impact decisions on the strategy. The cryo-EM flowchart of PSII in Figure 4.1 illustrates the stages involved in structure determination. In this chapter, our recent 3D model will be described and interpreted. The process of refining the data and testing various image processing software will be discussed. After collecting several cryo-EM data sets, the strategy for obtaining an optimal data set has been determined, though improvements can be made for future cryo-EM experiments.

Our most recent model at 17 Å resolution, elucidates the overall size, shape, and dimensions of the complex and the density can be compared to previously solved structures of PSII to gain an understand of the arrangements of subunits. The extrinsic subunits have an important biological role for the activity of PSII. Their locations on the complex can help further define their role in oxygen evolution activity. Identifying the positions of the extrinsic subunits can be performed by biochemically modifying PSII to remove specific extrinsic subunits. Obtaining a cryo-EM structure of a PSII complex with just PsbO bound, would elucidate the location of the subunits on the highly active C_2 complex or suggest a binding site for a second copy of PsbO. Addressed in Chapter 1, there is biochemical evidence of two copies of PsbO per monomer of plant PSII, though

current PSII structures have yet to support the experimental evidence [42-44]. The model shown in our data suggests an alternate binding of PsbO on the complex, but does not necessarily contradict the structures published in the past two years. We produce a highly active PSII sample, which does not undergo protein polishing steps like the previous strucutres. It is possible that structural changes occur when comparing a highly active PSII complex to the inactive structures. The extrinsic subunits have a key role in activity due to their close relationship with the oxygen evolving center (OEC). If such a relationship exists in plant PSII, then a structure of an active complex would show conformational changes when compared to a structure that is inactive.

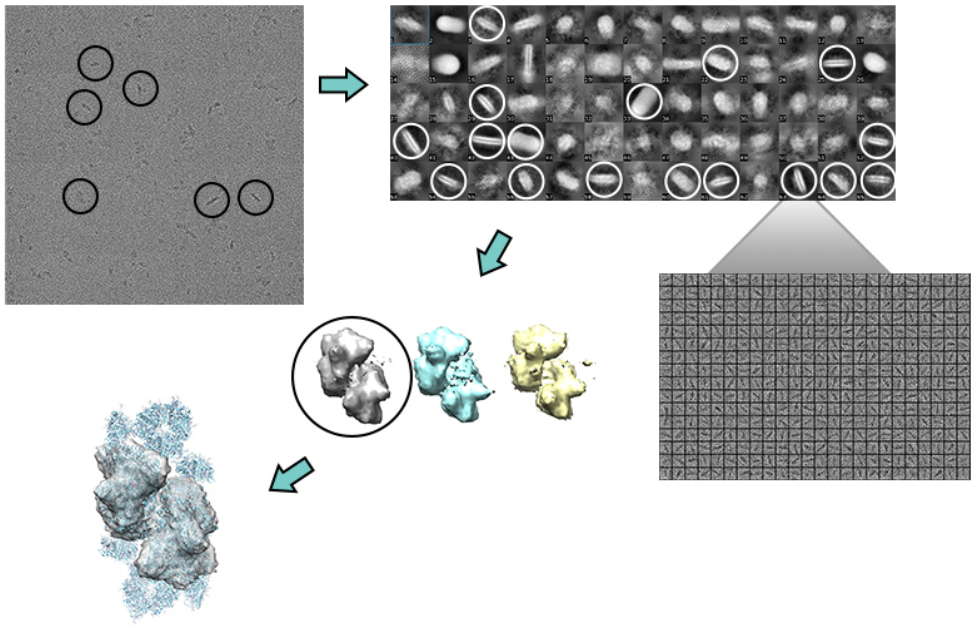


Figure 4.1 Diagram of PSII SPA cryo-EM describing the steps involved toward obtaining a 3D reconstruction. Particle images of PSII projections are selected and 2D averaging performed. Each 2D class represents many identical particle images that have been aligned and averaged. The optimal 2D particle classes are selected for 3D reconstruction and the best 3D class is refined and interpreted.

4.2 Results

4.2.1 Image processing and 3D reconstruction

Cryo-EM data collection was collected on the Titan Krios at Florida State University (FSU) via an NIH U24 consortium grant. The data sets were collected automatically using Leginon software. Images were selected based on qualities such as ice conditions, particle numbers, and contaminating artifacts. Automated and manual particle picking algorithms are part of the Appion software package at FSU, in addition to frame alignment and CTF estimation [98]. Due to the quality of our data, manual picking was the best option. Automated particle picking methods were attempted but did not result in meaningful 2D averages. Once the particles were picked, a particle stack was generated. Migration of this data into separate software packages such as Relion, EMAN, FREALIGN, cryoSPARC, or cisTEM allowed us to move directly to the 2D averaging and 3D reconstruction steps. Eliminating particles that did not correspond to the C_2 PSII enabled us to proceed to 3D reconstruction using an initial model derived from the PDB: 3JCU structure. To avoid model bias, the initial model resolution limit was set to 30 Å or greater. In addition, the model discussed in this work was selected based on several iterations of 3D refinement. Of the approximately 53,000 particles, 14% were used to generate a C_2 PSII reconstruction.

4.2.2 Cryo-EM 2D class averages offer insights on particle distribution

The 2D averages provided important information on the types of particles that existed in our sample purification (Chapter 3, Figure 3.13). For example, evidence indicated that there were two types of PSII particles within the data set (Figure 4.2). One

type was the C_2 core PSII dimer, and the other type may have been the C_2S_2 or perhaps two C_2 cores attached end-to-end (Figure 4.2 C and D). The particles of interest were approximately 20 nm in length and 10 nm in width. Some of the particles picked were longer than 20 nm, but these classes were not included for further processing. Additionally, some averages were ice contamination, non-optimal particles, or particles too close to the edge of an image, which were also removed before beginning 3D classification. The 2D class averages indicated that our data consists predominantly of side-view orientations (perpendicular to the place of the membrane) of PSII.

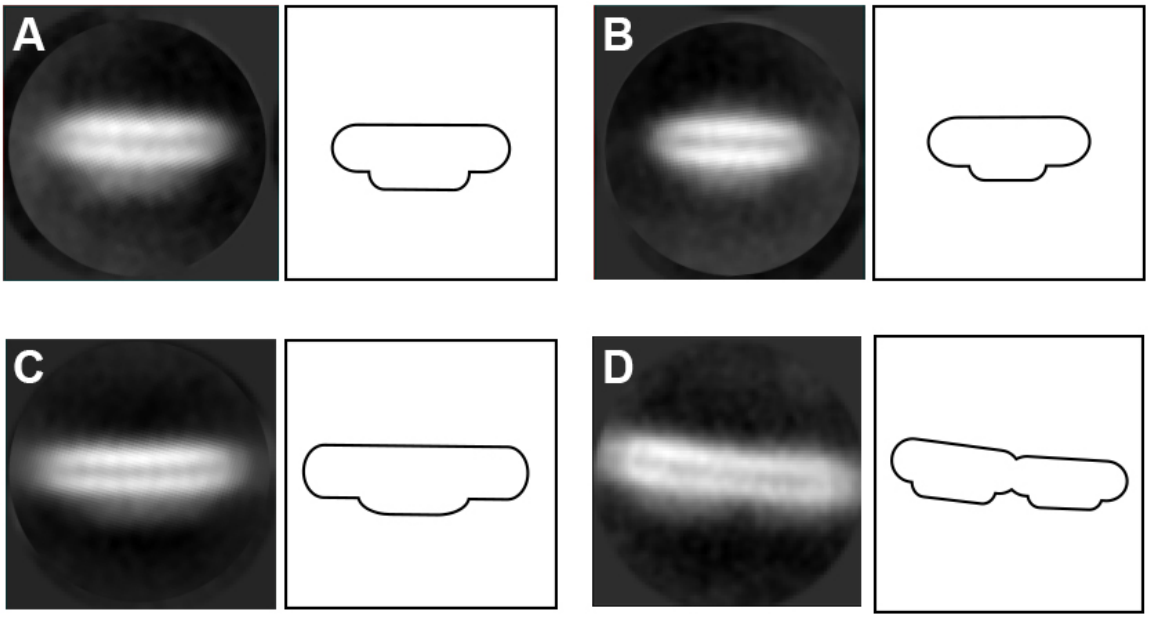


Figure 4.2 Cryo-EM 2D class averages of OPDM PSII. A select few classes generated by reference-free 2D classification in cisTEM showed possible PSII arrangements and orientations within this purification. Schematic representations are depicted to the right of each average to aid in the explanation of our interpretation. (A) C_2-type particle side-view. (B) C_2-type particle side-view, perpendicular to (A). (C) Potential side-view of a C_2S_2-type particle. (D Two end-to-end C_2-type particles. All box sizes represent 320 Å.

4.2.3 Overall architecture of PSII

Initially two data sets were collected of the same OPDM PSII sample. The first data set of C_2 PSII (Figure 4.3 A & 4.4) was collected using Quantifoil grids. These unique TEM grids are made of copper mesh with perforated carbon surfaces, with each hole being 2µm in diameter and spaced 2µm apart. The protein particles were suspended in ice, but had low concentrations and low contrast. The second, most recent data set (Figure 4.5, 4.6 & 4.7) was collected using C-flat perforated carbon grids (similar to Quantifoil), but these grids had a thin, continuous carbon layer underneath the holes, which served as a "well" to improve particle concentration. The structure and interpretation of the second C_2 PSII data set will be the predominant focus of this chapter.

The cryo-EM reconstruction represented the highly active C_2-type PSII core complex in *Spinacia oleracea*. Through biochemical characterization, it was confirmed that the sample was highly active and contained the important extrinsic subunits involved in oxygen evolution (Chapter 2, Results). The PSII complex forms a homodimer with two-fold symmetry running diagonally across the membrane (Figure 4.5). Using UCSF Chimera's rigid-body map fitting, our cryo-EM model was compared to the PDB structures for *T. vulcanus* PSII and *S. oleracea* PSII, PDB 3WU2 and 3JCU, respectively. The results show overall similar size, subunit arrangement, and geometry (Figure 4.6). The cyanobacterial structure (green) fit with our model with approximately 80% agreement. The spinach structure, which our initial model was derived from, matched to our model with 88% agreement when aligning the C_2 cores together.

The overall dimensions of the C_2 homodimer complex are 240 Å in length, 150 Å in width, and 160 Å in height, which corresponds closely to previous structures of inactive complexes (Figure 4.5 & 4.6). Each monomer of PSII contains the four major intrinsic subunits, CP43 (purple), CP47 (teal), D1 (red), and D2 (blue), along with some of the low molecular weight intrinsic polypeptides (PsbE, F, H, L, R, X, Y). In Figure 4.5, the extrinsic subunits PsbO (brown), PsbP (yellow), and PsbQ (cyan) are shown in their approximate locations forming a crown-shaped heterotrimer on the luminal side of the complex. $PsbT_n$ was also found in the mass spectrometry data (Chapter 2, Table 2.5), but was not illustrated on the model. The binding sites of the three, major extrinsic subunits roughly compare to the locations determined from the cyanobacteria and spinach PSII structures (Figure 4.6), however an alternate theory for the positions of PsbO and PsbP are shown in Figure 4.5. In Figure 4.4, the extrinsic subunits are illustrated based on the cyanobacterial and plant PSII structures.

An arrangement of the extrinsic subunits (PsbO, PsbP and PsbQ), where PsbO and PsbP are in opposite positions (Figure 4.3 B & C, and Figure 4.5), may be possible based on previous TEM experiments where the extrinsic subunits were removed [42, 45, 46, 164]. In Figure 4.7, our cryo-EM model was compared to the 3JCU spinach cryo-EM structure, which was rendered to 17 Å resolution. The overlay depicts densities that do not align, which could indicate differences between the active versus inactive structures. A higher resolution structure or a structure lacking PsbP and PsbQ (Figure 4.3 D) will be able to confirm the position of PsbO in an active PSII complex.

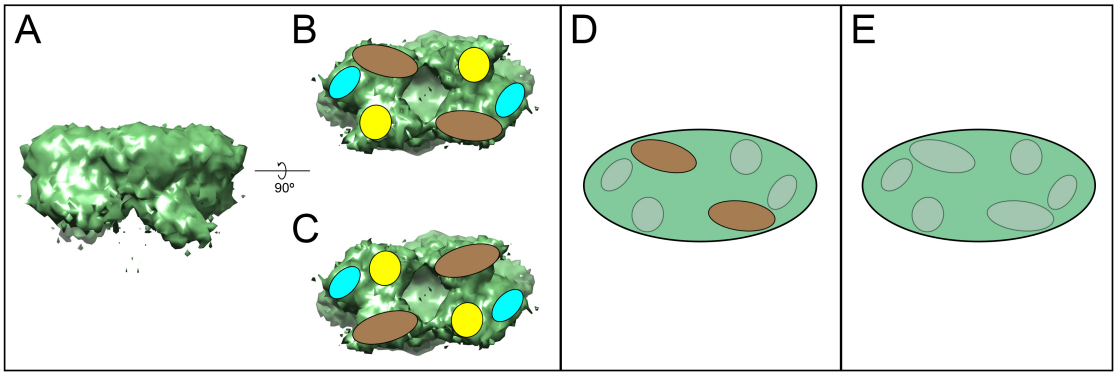


Figure 4.3 Cryo-EM model of PSII and proposed models of the complex under different conditions. (A) depicts our preliminary cryo-EM model of C_2 PSII in its side view with extrinsic subunits facing down. (B) and (C) illustrate the two differing arrangements of the extrinsic subunits, PsbO (brown), PsbP (yellow) and PsbQ (cyan). After salt-wash treatment (D), the expected structure of C_2 PSII would have just PsbO, and after urea-wash treatment (E), all extrinsic subunits would be removed.

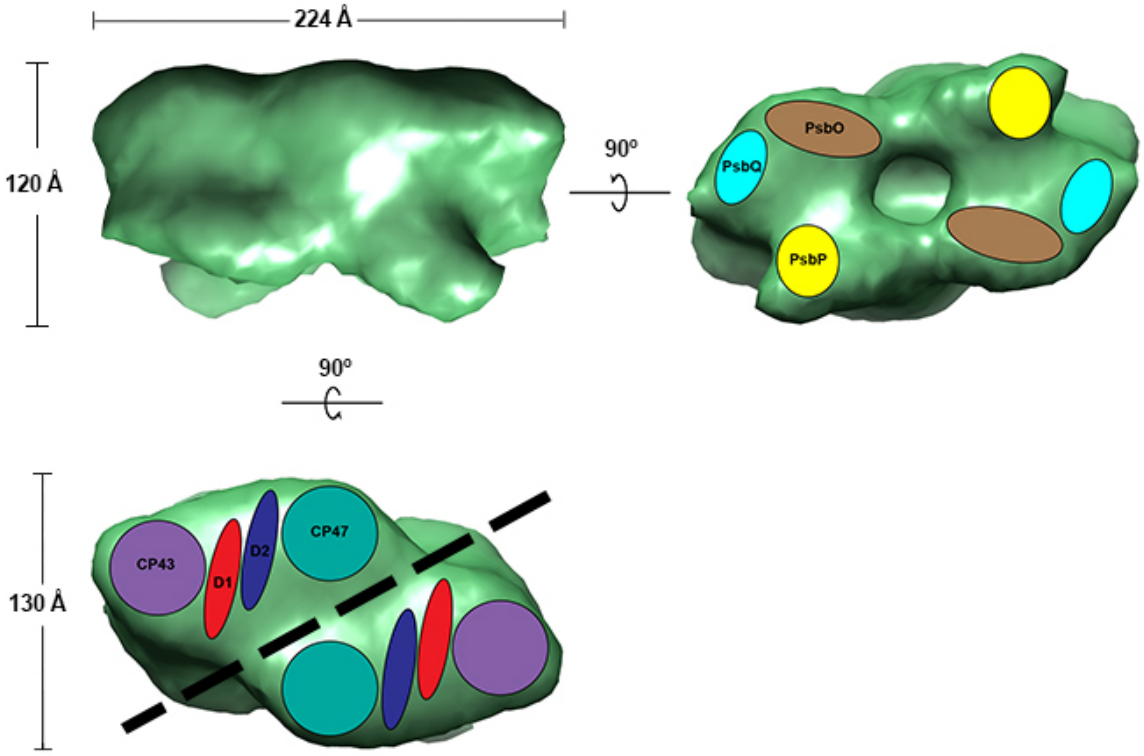


Figure 4.4 A first cryo-EM model of the highly active photosystem II C_2 core complex calculated from data collected using Quantifoil grids without carbon backing. In the upper left, the model is shown in its side view with extrinsic subunits facing down. In the upper right model, the lumenal face of the complex is shown with extrinsic subunits PsbO (brown), PsbP (yellow), and PsbQ (cyan). The model in the lower left corresponds to the stromal face of the complex with CP43 (purple), CP47 (teal), D1 (red), and D2 (blue). The line of symmetry is shown in black dotted lines, which separate each monomer of PSII.

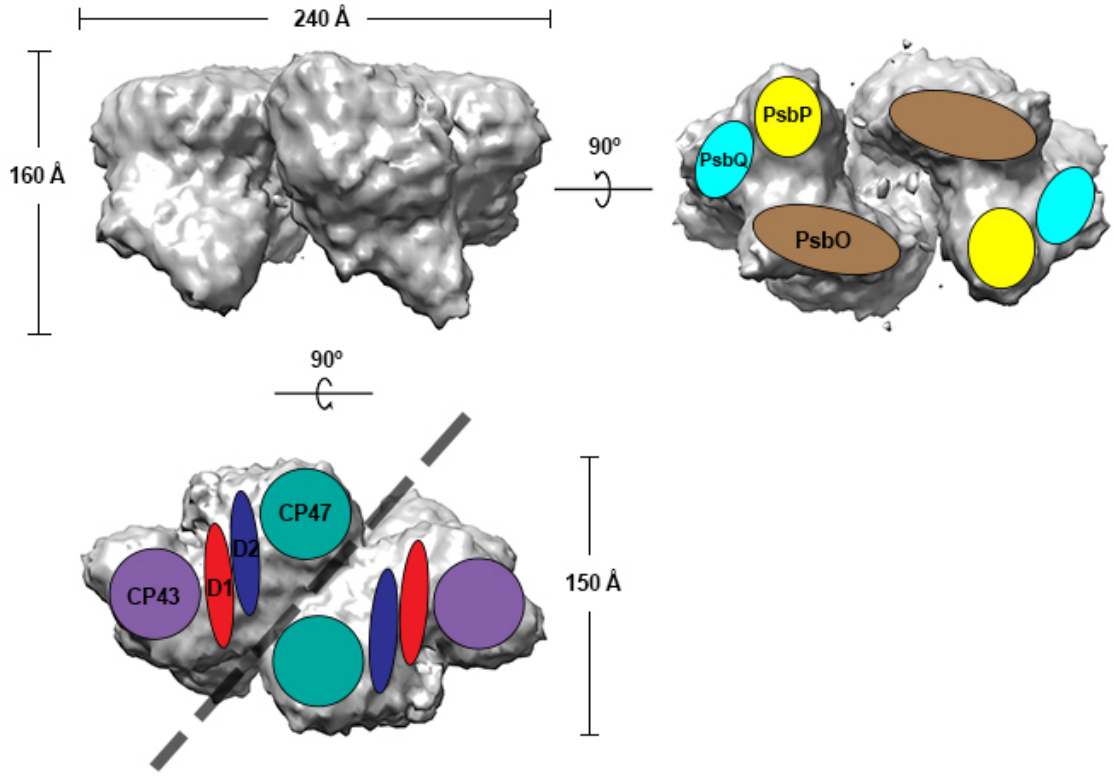


Figure 4.5 Current cryo-EM model of the highly active photosystem II C_2 core complex calculated from data collected using C-Flat grids with carboon backing. In the upper left, the model is shown in its side view with extrinsic subunits facing down. In the upper right model, the lumenal face of the complex is shown with extrinsic subunits PsbO (brown), PsbP (yellow), and PsbQ (cyan). The model in the lower left corresponds to the stromal face of the complex with CP43 (purple), CP47 (teal), D1 (red), and D2 (blue). The line of symmetry is shown in black dotted lines, which separate each monomer of PSII.

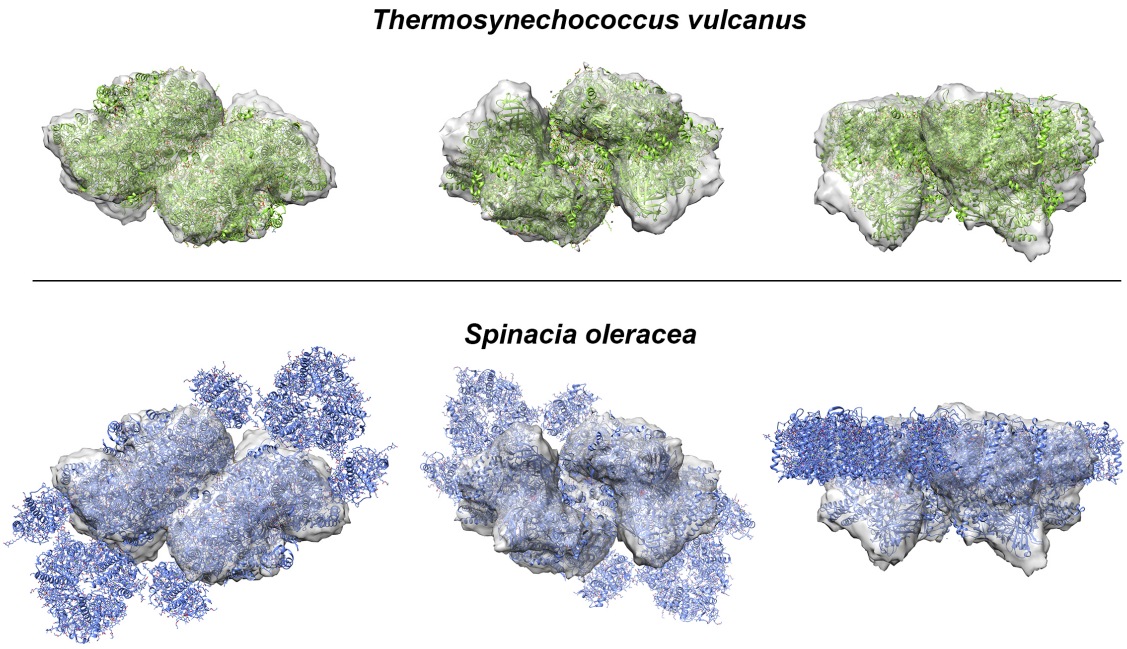


Figure 4.6 Comparison of the highly active cryo-EM model with previously published PSII structures. Our cryo-EM model was fit with a crystal structure of the *T. vulcanus* (PDB: 3WU2, green) and the cryo-EM structure of *S. oleracea* (PDB: 3JCU, blue). The structures had approximately 80% and 88% agreement for *T. vulcanus* and *S. oleracea*, respectively.

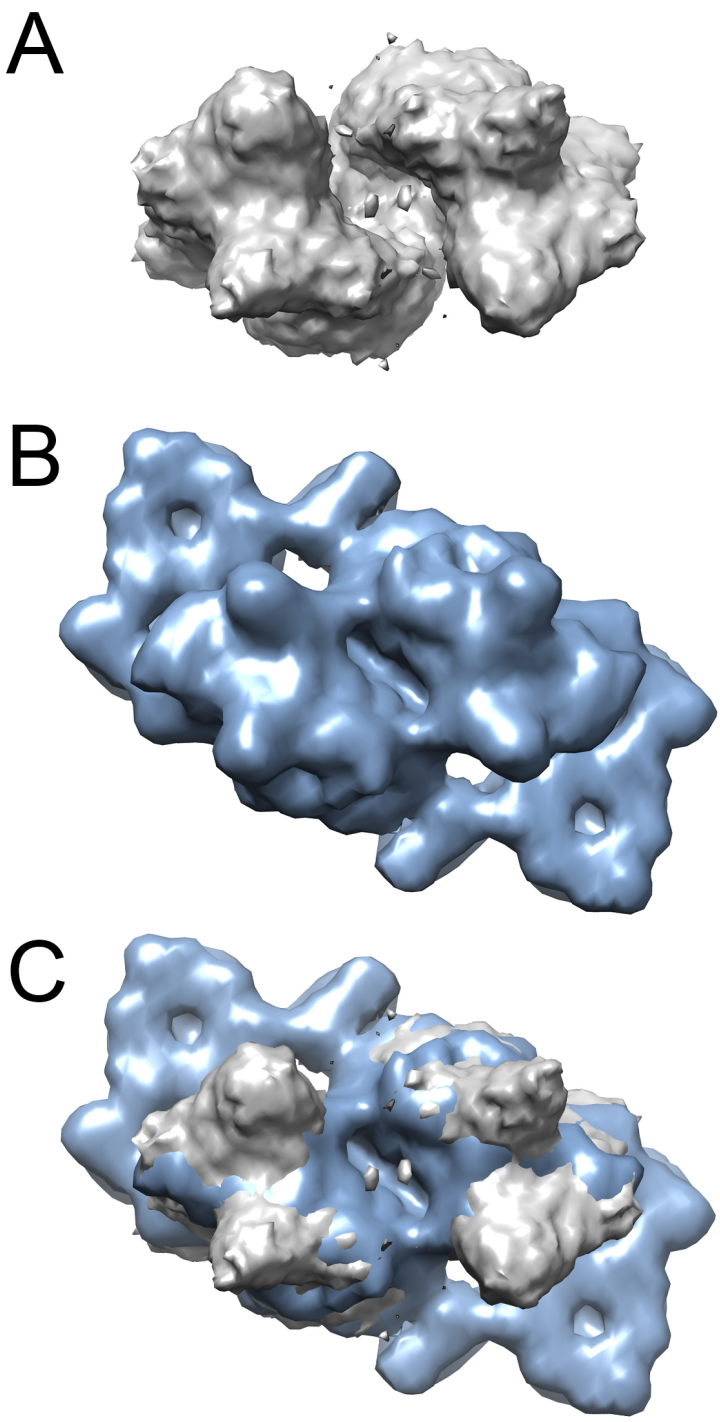


Figure 4.7 Comparison of the highly active PSII versus the inactive PSII cryo-EM models. (A) shows the highly active C_2 PSII (grey). (B) shows the C_2S_2 PSII (PDB: 3JCU) rendered at 17 Å resolution for comparison. (C) shows an overlay of the two models.

4.3 Discussion

Previous cryo-EM studies of the inactive PSII complex were at chlorophyll concentrations of 3 mg/mL [18]. However, because our preparation consists of predominantly C_2 cores (Chapter 3, Figure 3.13), the relative abundance of protein to chlorophyll is much higher at 3 mg/mL (Chapter 2, Discussion). Loading 3 mg/mL of our sample is not equivalent to loading 3 mg/mL of the sucrose gradient C_2S_2 sample. The viscosity of the C_2 PSII sample at chlorophyll concentrations higher than 2 mg/mL was too high for cryo-EM grid preparation. Using a lower concentration of 1 mg/mL and collecting a larger number of images, it was possible to obtain a first 3D model of the highly active PSII.

It is likely that the model we present contains the major intrinsic and extrinsic subunits of PSII. This conclusion can be made based on the biochemical results such as the SDS-PAGE, oxygen evolution assay, and mass spectroscopy analysis (Chapter 2). The three large densities on the lumenal facing side of the model suggest the presence of the major extrinsic subunits, PsbO, PsbP and PsbQ (Figure 4.7 A). The twelve low-molecular weight, intrinsic polypeptides are likely present in some combination within our complex. Though the SDS-PAGE data does not resolve these small subunits, it is known from prior literature that these proteins facilitate the dimerization between the larger subunits of PSII [18] and our cryo-EM model clearly shows a homodimer (Figure 4.5). Mass spectrometry analysis showed small intrinsic subunits in our preparation, such as PsbH, PsbL, PsbX and PsbY. Most of the other low molecular weight, intrinsic subunits were under 150 amino acids, so it is possible that some were undetected via mass spectrometry since small changes in mass could alter the mass to charge ratio and

result in a peptide that cannot be identified from the given database. However, the absence of these smaller proteins would likely lead to the complex disassembling, which is not seen in TEM. It is possible that some of the small intrinsic proteins that coat the outer perimeter of the C_2 PSII, where antenna complexes are bound, could be missing (subunits J, K, X, and Z), which could suggest why the antennae proteins are not bound to our complex. The subunits that exist at the monomer-monomer interface are more crucial for dimerization and are likely to be present (subunits I, L, M, T_c, W).

The placement of PsbO and PsbP in our model is currently under investigation. The literature from the early 2000s experimented on the removal of the extrinsic subunits, and they all agree with the placement of PsbO and PsbP, as described in Figure 4.3 C and Figure 4.5 [42, 45, 46, 164]. In the recent structures of PSII from 2016 and 2017, it was mentioned by some groups that the extrinsic subunit density was weaker due to partial loss of the extrinsic subunits, which agrees with their lower oxygen evolution rates [18, 20]. Cyanobacteria crystal structures and individually crystallized extrinsic subunits were used during model building for these recent structures, but cyanobacteria do not have PsbP and PsbQ. They also only have one copy of PsbO, where plants have two copies. Our strategy for cryo-EM provides an important foundation for future experiments to elucidate a salt-washed PSII structure, where PsbP and PsbQ have been removed. This would isolate PsbO, reveal its location on the complex, and give key insights to whether there might be a second copy of PsbO on a highly active PSII. This data could potentially contradict the current structures of plant PSII.

In Figure 4.7 C, one can visualize the differences in the extrinsic densities. This could correspond to shifts in the positions of the subunits in the active versus inactive

complexes. These shifts correspond to the extrinsic densities of PsbO, PsbP and PsbQ. One possibility that explains the shifts in density is the fact that our sample is more active than the structure it was compared to. The extrinsic subunits are known to play a key role in the activity of plant PSII, so it is reasonable to conclude that a loss or reduction of oxygen evolving activity results in a conformational change of the extrinsic subunits.

The 2D class averages contained particles that were longer than the typical 20 nm length C_2 particle (Figure 4.2 C). These particles were removed from the data before continuing to 3D reconstruction. However, if there are enough of these larger particles, it is possible to obtain a reconstruction of the C_2S_2 type PSII. This will be discussed more in the future directions in Chapter 5.

4.4 Experimental Procedures

4.4.1 Data collection, image processing, classification and reconstruction

The initial micrographs were screened for the best images using Appion [98]. Frame alignment was performed via MOTIONCORR, CTF estimation via CTFFIND4, and manual picking (Appion manual picker) was also performed using this software suite [98, 128]. Particle stacks were generated and exported from Florida State's High Performance Computing center to Georgia Tech's Partnership for Advanced Computing Environment (PACE) for further image processing via cisTEM [132]. Reference-free 2D classification was performed to generate a series of class averages. Approximately 20 classes (~54,00 particles) corresponding to the C_2 core PSII were selected for 3D classification. An initial model was generated from the PDB 3JCU structure using UCSF Chimera [165]. The antenna proteins (LHCII, CP26, CP29) were removed, leaving just the core C_2 complex. The initial model limit was set to 50 Å to remove potential bias in the starting reference. Two-fold (C2) symmetry was applied in 3D refinement. After 3D reconstruction and refinement, a 17 Å-resolution map was calculated to identify the location of the extrinsic subunits..

4.4.2 Model interpretation, measurements, and comparisons

The model was rendered in UCSF Chimera and measurements, rigid-body comparative analysis, and figures were generated with the software [165]. Due to the data being binned by two, the Å/pixel value was 2.02. Therefore, the model was scaled to a voxel size of 3.67 to compare with the PDB structures of 3JCU and 3WU2.

CHAPTER 5. CONCLUSIONS AND FUTURE DIRECTIONS

5.1 General considerations for cryo-EM studies of photosystem II

Photosystem II is a large, multi-subunit membrane protein that catalyzes the oxidation of water and creates O_2. Purification of this membrane protein is not trivial. PSII can have some degree of conformational heterogeneity, various flexible states and sizes of the complex, as well as preferred orientations. When considering the best method of solubilizing membrane proteins, sometimes amphipole-lipid scaffolds, such as nanodiscs, can be considered [166, 167]. Due to the robust preparation of our PSII sample and the high activity associated with it, we did not explore the use of these amphipole molecules for solubilization of PSII. In this work, the testing of detergents resulted in a sample suitable for cryo-EM. The strategy included retention of the high oxygen evolving activity of the sample. Key factors such as exchanging the detergent from OTG to DDM and eliminating a protein polishing step (sucrose gradient centrifugation) were critical in maintaining optimal activity and good contrast for cryo-EM.

5.1.1 Purification PSII

Ideally if one could control the growth and harvest of the spinach leaves, it would be possible to regulate potential modifications, mutations, environmental stress, differing strains, and other concerns that might be associated with isolating the protein from market spinach. However, at our current resolution, the minor changes that might be seen from factors such as these would be negligible. The purification of highly active PSII is

not a new method, but some alterations of the original OTG-PSII preparation were necessary in order to study this sample by cryo-EM. Biochemical assays were coupled with any new changes to the grid preparation in order to ensure the high oxygen evolving activity was not lost.

This purification method is popular among biochemists that study PSII by various spectroscopic methods. It is important to understand how the structure varies in an active versus inactive PSII complex. SPA cryo-EM has a major advantage because some heterogeneity can be present in a protein sample, and one can perform a targeted selection of the particles of interest during the image processing step.

5.1.2 Biochemical analysis

Several methods of monitoring the activity and integrity of PSII were used in this work. SDS-PAGE allowed us to identify the intrinsic and extrinsic subunits critical for oxygen evolving activity. These results revealed that our preparation of PSII contained the major intrinsic subunits, CP43, CP47, D1, and D2. It was also helpful in identifying the presence of the major extrinsic subunits, PsbO, PsbP, and PsbQ. In order to answer questions about the role of the extrinsic subunits, we performed salt wash and urea wash treatments on the sample. The SDS-PAGE results showed that PsbP and PsbQ could be removed with a salt wash, and all the extrinsic subunits could be removed with a urea wash (Chapter 3, Figure 3.7). Some proteins, however, do not appear in a Coomassie stained gel. Therefore, it was important to perform mass spectrometry of the sample to identify the presence of other subunits and confirm the major subunits of the C_2 PSII complex. These results confirmed that our sample contains all the major intrinsic and

extrinsic subunits, and we were also able to identify peptides from PSII subunits that were undetected by SDS-PAGE (PsbH, PsbL, PsbTn, PsbX, PsbY) (Chapter 2, Table 2.5).

PSII is a unique membrane protein because it contains pigments which can absorb light and give unique signatures that indicate the absence or presence of certain antenna proteins. Because the antenna proteins of PSII, which contain chlorophyll *b*, are absent in the C_2 PSII complex, UV-vis absorbance spectroscopy was a conclusive method of determining the reduction of chlorophyll *b* absorbance in the OTG-PSII and OPDM samples. These results indicated that the LHCII, CP29, CP26, and CP24 antenna proteins were reduced in the OTG-PSII and the OPDM sample (Chapter 2, Figure 2.6).

The final biochemical assay performed, which was the most direct method of validating our sample's activity, was the oxygen evolution assays. By creating an artificial environment for the PSII sample to oxidize water, the amount of oxygen produced could be quantified. Our results indicated high levels of oxygen evolving activity for the OPDM samples. This method is standardized and is how we can conclude that our protein sample is highly active compared to previous PSII samples (Chapter 2, Table 2.3).

5.1.3 Cryo-EM of PSII

The detergent exchange from OTG to DDM was a critical step for this project because it allowed us to identify C_2 core particles and conclude that they were the predominant type of particle in our sample. Several other important factors had to be considered for cryo-EM. The PSII sample was found to be too viscous above 2 mg/mL of

chlorophyll. Below 1 mg/mL however, the sample was too dilute that it resulted in too few particles per image. Two grid preparation methods were employed for our data sets. The first method uses the perforated carbon film (Quantifoil), which allowed the protein to adopt more orientations. The second method incorporated a carbon-backed grid (C-flat), which created carbon wells for the protein to adhere to a carbon surface. This method improved particle concentration, but it may somewhat limit the resolution due to the thin layer of carbon as opposed to vitreous ice. Conditions such as blot time, humidity within the Vitrobot, blot force, and glow discharging grids were considered when optimizing the cryo-EM conditions.

The results from cisTEM showed 2D averages that clearly indicated the presence of extrinsic subunits in the PSII particles. The 3D reconstruction showed the extrinsic subunits in shifted positions when compared to the inactive PSII structure. These differences may be attributed to the high activity of our PSII sample. The positioning of PsbO and PsbP was questioned due to past experiments that indicated one binding hypothesis [42, 45, 46, 164], and more recent data indicating a different binding arrangement (Chapter 4, Figure 4.3 B & C) [18-20, 39, 60]. We recently collected a salt-wash PSII data set of approximately 2700 images under identical cryo-EM conditions as the model presented in this work. In addition to processing the salt-wash data, improvements on the current model will result in a higher resolution structure of C_2 PSII, which could provide information about the location of PsbO, PsbP, and PsbQ.

5.2 Future work for the PSII project

There are many avenues of future exploration for this PSII project. Some advance the characterization of the sample and some improve the cryo-EM conditions for higher resolution. Starting with the biochemical experiments, the pigment content analysis of PSII can be performed by high performance liquid chromatography (HPLC,) which would confirm the abundance of chlorophyll molecules, β-carotene, lutein, neoxanthin, and violaxanthin present in the sample. Considering that our purification contains mostly the C_2 core PSII, but also has light harvesting antenna proteins, it can be predicted that the relative abundance of these molecules would be similar to what has been reported, however it would be important to know if any significant changes in the quantities of these pigments occur.

There are also lipids within and surrounding the complex that could be identified by tandem mass spectrometry. Galactolipids are the predominant expected lipid type in this sample. The work shown in this thesis reproduced PSII sample using a sucrose gradient centrifugation method, which resulted in a drop in oxygen evolving activity. However, by altering buffer conditions and pH, one could explore whether activity could be maintained.

The image processing strategy for the highly active PSII paves the way for future experiments, which will improve the resolution of this complex and allow for other types of PSII structures to be studied. The structures published previously report important high-resolution information, but their samples do not represent active PSII. Additionally,

the potential alternate localization of the extrinsic subunits was not discussed in these articles.

A current and ongoing experiment to study the changes in the PSII complex when extrinsic subunits are removed (either via salt wash or urea wash) is underway. The results of these experiments, specifically the salt washed PSII, will clarify the positions of the extrinsic subunits. This could help us determine alternate binding sites of the extrinsic subunits in an active PSII complex, or it could suggest the binding of a second copy of PsbO. Urea-washed PSII cryo-EM grids samples have been prepared and are ready for data collection. A cryo-EM structure of PSII with the major extrinsic subunits removed might result in overall conformational changes in the complex, which could help understand what occurs during subunit repair. These future experiments will aid in answering the biological question of the role of the extrinsic subunits in PSII. Continuing to refine the current PSII data set will improve the overall resolution and provide insights to the active complex as a whole.

Photoinhibition occurs when PSII is damaged by light. PSII subunit turnover is one of the many events that occur in the plant cell during this photo-damaging event. Protocols exist in the Barry lab to photoinhibit PSII, which would allow us to explore the structural changes at various time points during photoinhibition. This information could guide our understanding of the critical changes that occur in this process.

It was seen in our control OPDM sample that some potential C_2S_2 particles may exist in our preparation. During image processing, these particles were excluded, but if we separately process these larger PSII particles, we will obtain a C_2S_2 structure.

Additionally, we could investigate the salt wash and urea wash C_2S_2 PSII structures using the same conditions that worked for the C_2 PSII. These results would further improve our understanding of the differences between active and inactive PSII. In addition, it would aid in understanding the movement of light energy from the peripheral antenna proteins to the core of an active PSII complex.

The theme of these future studies revolve around specific modifications to the PSII sample followed by cryo-EM. The work of this thesis has provided a foundation for many new experiments to be explored under similar conditions that worked for our sample. By coupling biochemical characterizations with cryo-EM, the activity and integrity of PSII can be monitored, and the structural and functional details can be elucidated.

REFERENCES

1. Mojzsis, S.J., et al., *Evidence for life on Earth before 3,800 million years ago.* Nature, 1996. **384**(6604): p. 55.

2. Des Marais, D.J., *When did photosynthesis emerge on Earth?* Science, 2000. **289**(5485): p. 1703-1705.

3. Cardol, P. and A. Krieger-Liszkay, *From light capture to metabolic needs, oxygenic photosynthesis is an ever-expanding field of study in plants, algae and cyanobacteria.* Physiologia plantarum, 2017. **161**(1): p. 2-5.

4. Barber, J., *Photosystem II: the engine of life.* Quarterly reviews of biophysics, 2003. **36**(01): p. 71-89.

5. Andersson, B. and J.M. Anderson, *Lateral heterogeneity in the distribution of chlorophyll-protein complexes of the thylakoid membranes of spinach chloroplasts.* Biochimica et Biophysica Acta (BBA)-Bioenergetics, 1980. **593**(2): p. 427-440.

6. Koochak, H., et al., *The structural and functional domains of plant thylakoid membranes.* The Plant Journal, 2018.

7. Hill, R. and F. Bendall, *Function of the two cytochrome components in chloroplasts: a working hypothesis.* Nature, 1960. **186**(4719): p. 136.

8. Shen, J.-R., *The Structure of Photosystem II and the Mechanism of Water Oxidation in Photosynthesis.* Annual Review of Plant Biology, 2015. **66**(1): p. 23-48.

9. Diner, B.A. and F. Rappaport, *Structure, dynamics, and energetics of the primary photochemistry of photosystem II of oxygenic photosynthesis.* Annual review of plant biology, 2002. **53**(1): p. 551-580.

10. Van Gorkom, H., *Electron transfer in photosystem II.* Photosynthesis research, 1985. **6**(2): p. 97-112.

11. Rochaix, J.-D., *Reprint of: Regulation of photosynthetic electron transport.* Biochimica et Biophysica Acta (BBA)-Bioenergetics, 2011. **1807**(8): p. 878-886.

12. Kok, B., B. Forbush, and M. McGloin, *Cooperation of charges in photosynthetic O2 evolution–I. A linear four step mechanism.* Photochemistry and Photobiology, 1970. **11**(6): p. 457-475.

13. Joliot, P., G. Barbieri, and R. Chabaud, *Un nouveau modele des centres photochimiques du systeme II.* Photochemistry and Photobiology, 1969. **10**(5): p. 309-329.

14. Umena, Y., et al., *Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 A.* Nature, 2011. **473**(7345): p. 55-60.

15. Hankamer, B., et al., *Three-dimensional structure of the photosystem II core dimer of higher plants determined by electron microscopy.* Journal of structural biology, 2001. **135**(3): p. 262-269.

16. Rhee, K.-H., et al., *Two-dimensional structure of plant photosystem II at 8-Å resolution.* Nature, 1997. **389**(6650): p. 522.

17. Rhee, K.-H., et al., *Three-dimensional structure of the plant photosystem II reaction centre at 8[thinsp]A resolution.* Nature, 1998. **396**(6708): p. 283-286.

18. Wei, X., et al., *Structure of spinach photosystem II–LHCII supercomplex at 3.2 Å resolution.* Nature, 2016. **534**(7605): p. 69.

19. Su, X., et al., *Structure and assembly mechanism of plant C2S2M2-type PSII-LHCII supercomplex.* Science, 2017. **357**(6353): p. 815-820.

20. van Bezouwen, L.S., et al., *Subunit and chlorophyll organization of the plant photosystem II supercomplex.* Nature plants, 2017. **3**(7): p. 17080.

21. Mizusawa, N. and H. Wada, *The role of lipids in photosystem II.* Biochimica et Biophysica Acta (BBA)-Bioenergetics, 2012. **1817**(1): p. 194-208.

22. Gounaris, K., et al., *Structural and functional changes associated with heat-induced phase-separations of non-bilayer lipids in chloroplast thylakoid membranes.* FEBS letters, 1983. **153**(1): p. 47-52.

23. Sheng, X., et al., *Structural roles of lipid molecules in the assembly of plant PSII–LHCII supercomplex.* Biophysics Reports: p. 1-15.

24. Gounaris, K., D. Whitford, and J. Barber, *The effect of thylakoid lipids on an oxygen-evolving Photosystem II preparation.* FEBS Letters, 1983. **163**(2): p. 230-234.

25. Gross, J., *Chlorophylls*, in *Pigments in Vegetables*. 1991, Springer. p. 3-74.

26. Papageorgiou, G.C., *Chlorophyll a fluorescence: a signature of photosynthesis*. Vol. 19. 2007: Springer Science & Business Media.

27. Shen, G., J.J. Eaton-Rye, and W.F. Vermaas, *Mutation of histidine residues in CP47 leads to destabilization of the photosystem II complex and to impairment of light energy transfer.* Biochemistry, 1993. **32**(19): p. 5109-5115.

28. Bricker, T.M. and L.K. Frankel, *The structure and function of CP47 and CP43 in Photosystem II.* Photosynthesis Research, 2002. **72**(2): p. 131.

29. Alfonso, M., et al., *Core antenna complexes, CP43 and CP47, of higher plant photosystem II. Spectral properties, pigment stoichiometry, and amino acid composition.* Biochemistry, 1994. **33**(34): p. 10494-10500.

30. Trebst, A., *The three-dimensional structure of the herbicide binding niche on the reaction center polypeptides of photosystem II.* Zeitschrift für Naturforschung C, 1987. **42**(6): p. 742-750.

31. Allen, J., et al., *Structure of the reaction center from Rhodobacter sphaeroides R-26: the protein subunits.* Proceedings of the National Academy of Sciences, 1987. **84**(17): p. 6162-6166.

32. De Las Rivas, J., B. Andersson, and J. Barber, *Two sites of primary degradation of the D1-protein induced by acceptor or donor side photo-inhibition in photosystem II core complexes.* FEBS letters, 1992. **301**(3): p. 246-252.

33. Murata, N., et al., *Photoinhibition of photosystem II under environmental stress.* Biochimica et Biophysica Acta (BBA)-Bioenergetics, 2007. **1767**(6): p. 414-421.

34. Aro, E.-M., I. Virgin, and B. Andersson, *Photoinhibition of photosystem II. Inactivation, protein damage and turnover.* Biochimica et Biophysica Acta (BBA)-Bioenergetics, 1993. **1143**(2): p. 113-134.

35. Shipton, C.A. and J. Barber, *Photoinduced degradation of the D1 polypeptide in isolated reaction centers of photosystem II: evidence for an autoproteolytic process triggered by the oxidizing side of the photosystem.* Proceedings of the National Academy of Sciences, 1991. **88**(15): p. 6691-6695.

36. McEvoy, J.P. and G.W. Brudvig, *Water-splitting chemistry of photosystem II.* Chemical reviews, 2006. **106**(11): p. 4455-4483.

37. Cox, N., et al., *Electronic structure of the oxygen-evolving complex in photosystem II prior to OO bond formation.* Science, 2014. **345**(6198): p. 804-808.

38. Murata, N. and M. Miyao, *Extrinsic membrane proteins in the photosynthetic oxygen-evolving complex.* Trends in Biochemical Sciences, 1985. **10**(3): p. 122-124.

39. Suga, M., et al., *Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses.* Nature, 2015. **517**(7532): p. 99.

40. Yi, X., et al., *The manganese-stabilizing protein is required for photosystem II assembly/stability and photoautotrophy in higher plants.* Journal of Biological Chemistry, 2005. **280**(16): p. 16170-16174.

41. Offenbacher, A.R., B.C. Polander, and B.A. Barry, *An intrinsically disordered photosystem II subunit, PsbO, provides a structural template and a sensor of the*

hydrogen-bonding network in photosynthetic water oxidation. Journal of Biological Chemistry, 2013. **288**(40): p. 29056-29068.

42. Nield, J., et al., *Three-dimensional electron cryo-microscopy study of the extrinsic domains of the oxygen-evolving complex of spinach assignment of the PsbO protein.* Journal of Biological Chemistry, 2002. **277**(17): p. 15006-15012.

43. Popelkova, H., et al., *Inorganic cofactor stabilization and retention: the unique functions of the two PsbO subunits of eukaryotic photosystem II.* Biochemistry, 2008. **47**(47): p. 12593-12600.

44. Xu, Q. and T. Bricker, *Structural organization of proteins on the oxidizing side of photosystem II. Two molecules of the 33-kDa manganese-stabilizing proteins per reaction center.* Journal of Biological Chemistry, 1992. **267**(36): p. 25816-25821.

45. Boekema, E.J., et al., *Localization of the 23-kDa subunit of the oxygen-evolving complex of photosystem II by electron microscopy.* European journal of biochemistry, 1998. **252**(2): p. 268-276.

46. Boekema, E.J., et al., *Conformational changes in photosystem II supercomplexes upon removal of extrinsic subunits.* Biochemistry, 2000. **39**(42): p. 12907-12915.

47. Hankamer, B., J. Barber, and E.J. Boekema, *Structure and membrane organisation of photosystem II in green plants.* Annual Review of Plant Physiology and Plant Molecular Biology, 1997. **48**(1): p. 641-671.

48. MacColl, R., *Cyanobacterial Phycobilisomes.* Journal of Structural Biology, 1998. **124**(2): p. 311-334.

49. Satoh, S., et al., *Chlorophyll b expressed in cyanobacteria functions as a light-harvesting antenna in photosystem I through flexibility of the proteins.* Journal of Biological Chemistry, 2001. **276**(6): p. 4293-4297.

50. Caffarri, S., et al., *Functional architecture of higher plant photosystem II supercomplexes.* The EMBO journal, 2009. **28**(19): p. 3052-3063.

51. Horton, P., A. Ruban, and R. Walters, *Regulation of light harvesting in green plants.* Annual review of plant biology, 1996. **47**(1): p. 655-684.

52. Santini, C., et al., *Three-dimensional structure of the higher-plant photosystem II reaction centre and evidence for its dimeric organization in vivo.* European journal of biochemistry, 1994. **221**(1): p. 307-315.

53. Enami, I., et al., *Structures and functions of the extrinsic proteins of photosystem II from different species.* Photosynthesis research, 2008. **98**(1-3): p. 349-363.

54. Shen, J.-R., M. Ikeuchi, and Y. Inoue, *Stoichiometric association of extrinsic cytochrome c 550 and 12 kDa protein with a highly purified oxygen-evolving*

photosystem II core complex from Synechococcus vulcanus. FEBS letters, 1992. **301**(2): p. 145-149.

55. Enami, I., et al., *Isolation and characterization of a Photosystem II complex from the red alga Cyanidium caldarium: association of cytochrome c-550 and a 12 kDa protein with the complex.* Biochimica et Biophysica Acta (BBA)-Bioenergetics, 1995. **1232**(3): p. 208-216.

56. Suzuki, T., et al., *Isolation and characterization of oxygen-evolving photosystem II complexes retaining the PsbO, P and Q proteins from Euglena gracilis.* Plant and cell physiology, 2004. **45**(9): p. 1168-1175.

57. Berthold, D.A., G.T. Babcock, and C.F. Yocum, *A highly resolved, oxygen-evolving photosystem II preparation from spinach thylakoid membranes: EPR and electron-transport properties.* FEBS Letters, 1981. **134**(2): p. 231-234.

58. Cardona, T., et al., *Charge separation in photosystem II: a comparative and evolutionary overview.* Biochimica et Biophysica Acta (BBA)-Bioenergetics, 2012. **1817**(1): p. 26-43.

59. Ballottari, M., et al., *Evolution and functional properties of Photosystem II light harvesting complexes in eukaryotes.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2012. **1817**(1): p. 143-157.

60. Albanese, P., et al., *Pea PSII-LHCII supercomplexes form pairs by making connections across the stromal gap.* Scientific reports, 2017. **7**(1): p. 10067.

61. Roose, J.L., et al., *The extrinsic proteins of photosystem II: update.* Planta, 2016. **243**(4): p. 889-908.

62. Yano, J., et al., *X-ray damage to the Mn4Ca complex in single crystals of photosystem II: a case study for metalloprotein crystallography.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(34): p. 12047-12052.

63. Grabolle, M., et al., *Rapid loss of structural motifs in the manganese complex of oxygenic photosynthesis by X-ray irradiation at 10–300 K.* Journal of Biological Chemistry, 2006. **281**(8): p. 4580-4588.

64. Kern, J., et al., *Room temperature femtosecond X-ray diffraction of photosystem II microcrystals.* Proceedings of the National Academy of Sciences, 2012. **109**(25): p. 9721-9726.

65. Boodhun, N., *Seeing is believing: structures and functions of biological molecules*. 2018, Future Science.

66. Carpenter, E.P., et al., *Overcoming the challenges of membrane protein crystallography.* Current opinion in structural biology, 2008. **18**(5): p. 581-586.

67. Tan, S., H.T. Tan, and M.C. Chung, *Membrane proteins and membrane proteomics.* Proteomics, 2008. **8**(19): p. 3924-32.

68. Anson, L., *Membrane protein biophysics.* Nature, 2009. **459**(7245): p. 343.

69. Hendrickson, W.A., *Atomic-level analysis of membrane-protein structure.* Nature structural & molecular biology, 2016. **23**(6): p. 464.

70. Lacapere, J.-J., et al., *Determining membrane protein structures: still a challenge!* Trends in biochemical sciences, 2007. **32**(6): p. 259-270.

71. Junge, F., et al., *Large-scale production of functional membrane proteins.* Cellular and Molecular Life Sciences, 2008. **65**(11): p. 1729-1755.

72. Lin, S.-H. and G. Guidotti, *Purification of membrane proteins*, in *Methods in enzymology*. 2009, Elsevier. p. 619-629.

73. Tate, C.G., *Practical considerations of membrane protein instability during purification and crystallisation*, in *Heterologous Expression of Membrane Proteins*. 2010, Springer. p. 187-203.

74. Dolder, M., A. Engel, and M. Zulauf, *The micelle to vesicle transition of lipids and detergents in the presence of a membrane protein: towards a rationale for 2D crystallization.* FEBS letters, 1996. **382**(1-2): p. 203-208.

75. Duquesne, K. and J.N. Sturgis, *Membrane protein solubilization*, in *Heterologous Expression of Membrane Proteins*. 2010, Springer. p. 205-217.

76. Rubinstein, J.L., *Structural analysis of membrane protein complexes by single particle electron microscopy.* Methods, 2007. **41**(4): p. 409-416.

77. Bai, X.-c., G. McMullan, and S.H.W. Scheres, *How cryo-EM is revolutionizing structural biology.* Trends in Biochemical Sciences, 2015. **40**(1): p. 49-57.

78. Doerr, A., *Structural biology: Cryo-EM goes high-resolution.* Nature methods, 2015. **12**(7): p. 598.

79. Ruska, E., *The early development of electron lenses and electron microscopy.* Microscopica acta. Supplement, 1980(Suppl 5): p. 1-140.

80. Lewis Jr, R., *Scattering of Electrons from Nuclear Charge Fluctuations.* Physical Review, 1956. **102**(2): p. 544.

81. Anderson, T.F., *Techniques for the preservation of three-dimensional structure in preparing specimens for the electron microscope.* Transactions of the New York Academy of Sciences, 1951. **13**(4 Series II): p. 130-134.

82. Crowther, R. and D. DeRosier, *The reconstruction of a three-dimensional structure from projections and its application to electron microscopy.* Proc. R. Soc. Lond. A, 1970. **317**(1530): p. 319-340.

83. Herman, G. and R. Lewitt, *Overview of image reconstruction from projections.* Image Reconstruction from Projections, 1979: p. 1-8.

84. Glaeser, R.M., *Electron crystallography of biological macromolecules*. 2007: Oxford University Press.

85. Zhu, D., et al., *Pushing the resolution limit by correcting the Ewald sphere effect in single-particle Cryo-EM reconstructions.* Nature Communications, 2018. **9**(1): p. 1552.

86. Crowther, R., *Procedures for three-dimensional reconstruction of spherical viruses by Fourier synthesis from electron micrographs.* Phil. Trans. R. Soc. Lond. B, 1971. **261**(837): p. 221-230.

87. Frank, J. and P. Penczek, *On the correction of the contrast transfer function in biological electron microscopy.* Optik, 1995. **98**(3): p. 125-129.

88. Sander, B., M.M. Golas, and H. Stark, *Automatic CTF correction for single particles based upon multivariate statistical analysis of individual power spectra.* Journal of Structural Biology, 2003. **142**(3): p. 392-401.

89. McEwen, B.F., K.H. Downing, and R.M. Glaeser, *The relevance of dose-fractionation in tomography of radiation-sensitive specimens.* Ultramicroscopy, 1995. **60**(3): p. 357-373.

90. Hegerl, R. and W. Hoppe, *Influence of electron noise on three-dimensional image reconstruction.* Zeitschrift für Naturforschung A, 1976. **31**(12): p. 1717-1721.

91. Baumeister, W., *Electron tomography: towards visualizing the molecular organization of the cytoplasm.* Current opinion in structural biology, 2002. **12**(5): p. 679-684.

92. Nannenga, B.L. and T. Gonen, *Protein structure determination by MicroED.* Current opinion in structural biology, 2014. **27**: p. 24-31.

93. Unwin, P.N.T. and R. Henderson, *Molecular structure determination by electron microscopy of unstained crystalline specimens.* Journal of molecular biology, 1975. **94**(3): p. 425-440.

94. Henderson, R. and P.N.T. Unwin, *Three-dimensional model of purple membrane obtained by electron microscopy.* Nature, 1975. **257**(5521): p. 28.

95. Radermacher, M., *Three-dimensional reconstruction of single particles from random and nonrandom tilt series.* Microscopy Research and Technique, 1988. **9**(4): p. 359-394.

96. Kühlbrandt, W., *Microscopy: cryo-EM enters a new era.* Elife, 2014. **3**: p. e03678.

97. Lu, P., et al., *Three-dimensional structure of human γ-secretase.* Nature, 2014. **512**(7513): p. 166.

98. Lander, G.C., et al., *Appion: an integrated, database-driven pipeline to facilitate EM image processing.* Journal of structural biology, 2009. **166**(1): p. 95-102.

99. de Haas, F., et al., *Comparative electron microscopy and image analysis of oxy- and deoxy-hemocyanin from the spiny lobster Panulirus interruptus.* Ultramicroscopy, 1993. **49**(1-4): p. 426-435.

100. Van Heel, M. and J. Frank, *Use of multivariates statistics in analysing the images of biological macromolecules.* Ultramicroscopy, 1981. **6**(1): p. 187-194.

101. Henderson, R., *The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules.* Quarterly reviews of biophysics, 1995. **28**(2): p. 171-193.

102. Merk, A., et al., *Breaking cryo-EM resolution barriers to facilitate drug discovery.* Cell, 2016. **165**(7): p. 1698-1707.

103. Glaeser, R.M. and K.A. Taylor, *Radiation damage relative to transmission electron microscopy of biological specimens at low temperature: a review.* Journal of Microscopy, 1978. **112**(1): p. 127-138.

104. Brenner, S. and R. Horne, *A negative staining method for high resolution electron microscopy of viruses.* Biochimica et biophysica acta, 1959. **34**: p. 103-110.

105. Kellenberger, E., M. Häner, and M. Wurtz, *The wrapping phenomenon in air-dried and negatively stained preparations.* Ultramicroscopy, 1982. **9**(1): p. 139-150.

106. Amos, L., R. Henderson, and P. Unwin, *Three-dimensional structure determination by electron microscopy of two-dimensional crystals.* Progress in biophysics and molecular biology, 1983. **39**: p. 183-231.

107. Dubochet, J., et al., *Cryo-electron microscopy of vitrified specimens.* Quarterly reviews of biophysics, 1988. **21**(2): p. 129-228.

108. Dubochet, J. and A. McDowall, *Vitrification of pure water for electron microscopy.* Journal of Microscopy, 1981. **124**(3): p. 3-4.

109. Thompson, R.F., et al., *An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology.* Methods, 2016. **100**: p. 3-15.

110. Taylor, K.A. and R.M. Glaeser, *Electron diffraction of frozen, hydrated protein crystals.* Science, 1974. **186**(4168): p. 1036-1037.

111. De Rosier, D. and A. Klug, *Reconstruction of three dimensional structures from electron micrographs.* Nature, 1968. **217**(5124): p. 130.

112. Cheng, Y., *Single-particle cryo-EM at crystallographic resolution.* Cell, 2015. **161**(3): p. 450-457.

113. Kühlbrandt, W., *The resolution revolution.* Science, 2014. **343**(6178): p. 1443-1444.

114. Bai, X.-C., G. McMullan, and S.H. Scheres, *How cryo-EM is revolutionizing structural biology.* Trends in biochemical sciences, 2015. **40**(1): p. 49-57.

115. Portmann, R. and T. Koller. *The divalent cation method for protein-free spreading of nucleic acid molecules*. in *Sixth European Congress Electron microscopy*. 1976.

116. Hayat, M.A., *Principles and techniques of electron microscopy. Biological applications*. 1981: Edward Arnold.

117. Fernandez-Leiro, R. and S.H. Scheres, *Unravelling biological macromolecules with cryo-electron microscopy.* Nature, 2016. **537**(7620): p. 339.

118. Dobro, M.J., et al., *Plunge freezing for electron cryomicroscopy*, in *Methods in enzymology*. 2010, Elsevier. p. 63-82.

119. Iancu, C.V., et al., *Electron cryotomography sample preparation using the Vitrobot.* Nature protocols, 2006. **1**(6): p. 2813.

120. Ranson, N.A., et al., *ATP-bound states of GroEL captured by cryo-electron microscopy.* Cell, 2001. **107**(7): p. 869-879.

121. Tan, Y.Z., et al., *Addressing preferred specimen orientation in single-particle cryo-EM through tilting.* Nature methods, 2017. **14**(8): p. 793.

122. Glaeser, R.M., *How good can cryo-EM become?* Nature methods, 2015. **13**(1): p. 28.

123. Saxton, W. and J. Frank, *Motif detection in quantum noise-limited electron micrographs by cross-correlation.* Ultramicroscopy, 1976. **2**: p. 219-227.

124. Suloway, C., et al., *Automated molecular microscopy: The new Leginon system.* Journal of Structural Biology, 2005. **151**(1): p. 41-60.

125. Fujiyoshi, Y., *The structural study of membrane proteins by electron crystallography.* Advances in biophysics, 1998. **35**: p. 25-80.

126. Avila-Sakar, A., et al., *Recording High-Resolution Images of Two-Dimensional Crystals of Membrane Proteins*, in *Electron Crystallography of Soluble and Membrane Proteins*. 2013, Springer. p. 129-152.

127. Clough, R., G. Moldovan, and A. Kirkland. *Direct detectors for electron microscopy*. in *Journal of Physics: Conference Series*. 2014. IOP Publishing.

128. Li, X., et al., *Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM.* Nature methods, 2013. **10**(6): p. 584.

129. McMullan, G., et al., *Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy.* Ultramicroscopy, 2014. **147**: p. 156-163.

130. Rawson, S., et al., *Methods to account for movement and flexibility in cryo-EM data processing.* Methods, 2016. **100**: p. 35-41.

131. Scheres, S.H., *RELION: implementation of a Bayesian approach to cryo-EM structure determination.* Journal of structural biology, 2012. **180**(3): p. 519-530.

132. Grant, T., A. Rohou, and N. Grigorieff, *cisTEM, user-friendly software for single-particle image processing.* Elife, 2018. **7**: p. e35383.

133. Frank, J., B. Shimkin, and H. Dowse, *SPIDER—a modular software system for electron image processing.* Ultramicroscopy, 1981. **6**(4): p. 343-357.

134. Tang, G., et al., *EMAN2: an extensible image processing suite for electron microscopy.* Journal of structural biology, 2007. **157**(1): p. 38-46.

135. Grigorieff, N., *FREALIGN: high-resolution refinement of single particle structures.* Journal of structural biology, 2007. **157**(1): p. 117-125.

136. Punjani, A., et al., *cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination.* Nature methods, 2017. **14**(3): p. 290.

137. van Heel, M., et al., *A new generation of the IMAGIC image processing system.* Journal of structural biology, 1996. **116**(1): p. 17-24.

138. Frank, J., et al., *SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields.* Journal of structural biology, 1996. **116**(1): p. 190-199.

139. Rath, B.K. and J. Frank, *Fast automatic particle picking from cryo-electron micrographs using a locally normalized cross-correlation function: a case study.* Journal of Structural Biology, 2004. **145**(1): p. 84-90.

140. Nicholson, W.V. and R.M. Glaeser, *Automatic particle detection in electron microscopy.* Journal of Structural Biology, 2001. **133**(2-3): p. 90-101.

141. Baldwin, P.R., et al., *Big data in cryoEM: automated collection, processing and accessibility of EM data.* Current opinion in microbiology, 2018. **43**: p. 1-8.

142. Ludtke, S.J., P.R. Baldwin, and W. Chiu, *EMAN: semiautomated software for high-resolution single-particle reconstructions.* Journal of structural biology, 1999. **128**(1): p. 82-97.

143. Radermacher, M., *Weighted back-projection methods*, in *Electron tomography*. 2007, Springer. p. 245-273.

144. Penczek, P.A., *Resolution measures in molecular electron microscopy.* Methods in enzymology, 2010. **482**: p. 73-100.

145. Frank, J., *Three-dimensional electron microscopy of macromolecular assemblies: visualization of biological molecules in their native state*. 2006: Oxford University Press.

146. Campbell, M.G., et al., *2.8 Å resolution reconstruction of the Thermoplasma acidophilum 20S proteasome using cryo-electron microscopy.* Elife, 2015. **4**: p. e06380.

147. Unser, M., B.L. Trus, and A.C. Steven, *A new resolution criterion based on spectral signal-to-noise ratios.* Ultramicroscopy, 1987. **23**(1): p. 39-51.

148. Liao, H.Y. and J. Frank, *Definition and estimation of resolution in single-particle reconstructions.* Structure (London, England : 1993), 2010. **18**(7): p. 768-775.

149. van Heel, M. and M. Schatz, *Fourier shell correlation threshold criteria.* Journal of Structural Biology, 2005. **151**(3): p. 250-262.

150. Barber, J., *Biological solar energy.* Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 2007. **365**(1853): p. 1007-1023.

151. Nogales, E., *The development of cryo-EM into a mainstream structural biology technique.* Nature methods, 2015. **13**(1): p. 24.

152. Mishra, R.K. and D.F. Ghanotakis, *Selective extraction of CP 26 and CP 29 proteins without affecting the binding of the extrinsic proteins (33, 23 and 17 kDa) and the DCMU sensitivity of a Photosystem II core complex.* Photosynthesis research, 1994. **42**(1): p. 37-42.

153. Hankamer, B., et al., *Isolation and Biochemical Characterization of Monomeric and Dimeric Photosystem II Complexes from Spinach and Their Relevance to the Organisation of Photosystem II In vivo.* European Journal of Biochemistry, 1997. **243**(1-2): p. 422-429.

154. Nield, J., et al., *3D map of the plant photosystem II supercomplex obtained by cryoelectron microscopy and single particle analysis.* Nature Structural & Molecular Biology, 2000. **7**(1): p. 44-47.

155. Ning, K., D. Fermin, and A.I. Nesvizhskii, *Comparative analysis of different label-free mass spectrometry based protein abundance estimates and their correlation with RNA-Seq gene expression data.* Journal of proteome research, 2012. **11**(4): p. 2261-2271.

156. Liu, H., R.G. Sadygov, and J.R. Yates, *A model for random sampling and estimation of relative protein abundance in shotgun proteomics.* Analytical chemistry, 2004. **76**(14): p. 4193-4201.

157. Ghanotakis, D.F., G.T. Babcock, and C.F. Yocum, *Calcium reconstitutes high rates of oxygen evolution in polypeptide depleted photosystem II preparations.* FEBS letters, 1984. **167**(1): p. 127-130.

158. Piccioni, R., G. Bellemare, and N.m.H. Chua, *Methods of polyacrylamide gel electrophoresis in the analysis and preparation of plant polypeptides.* Methods in chloroplast molecular biology/M. Edelman, RB Hallick, and NH Chua, editors, 1982.

159. Wiśniewski, J.R., et al., *Universal sample preparation method for proteome analysis.* Nature methods, 2009. **6**(5): p. 359.

160. Rinker, T.E., et al., *Microparticle-mediated sequestration of cell-secreted proteins to modulate chondrocytic differentiation.* Acta biomaterialia, 2018. **68**: p. 125-136.

161. Albanese, P., et al., *Isolation of novel PSII-LHCII megacomplexes from pea plants characterized by a combination of proteomics and electron microscopy.* Photosynthesis research, 2016. **130**(1-3): p. 19-31.

162. Uddin, Y.M. and I. Schmidt-Krey, *Inducing Two-Dimensional Crystallization of Membrane Proteins by Dialysis for Electron Crystallography*, in *Methods in enzymology*. 2015, Elsevier. p. 351-362.

163. Mindell, J.A. and N. Grigorieff, *Accurate determination of local defocus and specimen tilt in electron microscopy.* Journal of structural biology, 2003. **142**(3): p. 334-347.

164. Boekema, E.J., et al., *Supramolecular structure of the photosystem II complex from green plants and cyanobacteria.* Proceedings of the National Academy of Sciences, 1995. **92**(1): p. 175-179.

165. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis.* J Comput Chem, 2004. **25**(13): p. 1605-12.

166. Denisov, I.G. and S.G. Sligar, *Nanodiscs for structural and functional studies of membrane proteins.* Nature Structural and Molecular Biology, 2016. **23**(6): p. 481.

167. Hagn, F., et al., *Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins.* Journal of the American Chemical Society, 2013. **135**(5): p. 1919-1925.